

## Going national against the odds

*The British government is sensibly seeking to introduce a national curriculum for its schools, but may be defeated by the structure of the system it has inherited — and by some obstacles of its own making.*

WITH the possible exception of Japan, every industrialized country has taken fright at the task of providing its young people with a basic education. Two practical questions obtrude everywhere: how to decide what constitutes a preparation for life in the modern world, and how to recruit teachers capable of meeting the implied demands on them and their students. The first task is easier than the second.

Even when governments (as in the federal United States) say that it is not their job to specify a national curriculum, it usually emerges that there is a general understanding of the goals at which teachers should be aiming (in the United States, for example, the college boards tests have become a kind of *de facto* norm for the 30 per cent of young people staying on in higher education together with those who fail to do so). Much the same used to happen in Britain, where academic goals for schools used to be defined by public examinations organized by a series of regional Schools Examinations Boards, whose function until a few years ago was to certify the attainment of school-leavers at two ages — 16 and 18 years. But now, all that is to change: one of the least contentious and one of the few welcome proposals in the British government's Education Reform Bill is that there should be a national curriculum. Two documents published last week are part of the process of definition. The general case for this change is overwhelming. Over the years, the content of young people's education has become too dependent on whim and chance. The cumulative effect of a series of seeming 'reforms' stretching back for half a century has seriously eroded the content of what young people are taught at school. First, there has been a progressive relaxation of the matriculation (entrance) requirements of universities and other higher education institutions based on the assumption that it is absurd to require of intending students competence in fields unconnected with their future studies. Second, since the alarming discovery in the 1950s that the public examinations boards were foisting on the schools curricula better suited to a period half a century earlier, the content of what is taught to young people has increasingly become a matter for negotiation between interested groups of teachers and those academics in higher education who take an interest in these questions (who are few and, usually, undistinguished). Finally, the integrity of the curriculum in British schools has been undermined by the permissiveness that stems from the curious tradition of specialization that is the hallmark of British education, and by means of which students are encouraged and sometimes compelled to rehearse in their school careers the studies they will later follow at university (polytechnics and other institutions of higher education are less demanding.)

### Reservation

One important reservation is also telling. These complaints apply chiefly to the parts of Britain known as England and Wales (but the tradition in Northern Ireland is similar). They do not apply to Scotland (nor do the documents published last week), where specialization at school is less of a scandal and where discontents about the curriculum are, by comparison with England and Wales, unimportant. Is that not a sign that the time

has come, in Britain as a whole, simply to acknowledge that the Scots have been right all these years to insist that schools provide a general education and that university courses should last for four years? That, unremarkably, is the pattern followed almost everywhere else as well.

Meanwhile, although the new reforms fall short of true reform along those lines, there is to be a national curriculum. Last week's documents are proposals for the content of mathematics and science teaching in schools for students ranging in age from 5 to 16. The proposals deal with the whole of this age range as a whole, which is sensible. The group responsible for mathematics asks for more emphasis than in recent school practice on mental calculation and for an acknowledgement that calculators exist and will therefore be used, both of them sensible pragmatic adjustments to reality, but not the rejection of 'new' mathematics which it has been widely supposed to be. The mathematics panel is probably right to complain that 10 per cent of curriculum time is insufficient for its ambitious goals. For the science panel, the new dimension is that technology becomes a component of teaching even from the age of five, that divisions between different branches of science formally disappear (at least until the age of 16) and that the proportion of a student's time given over to study in this field should range between an eighth and a sixth at various ages.

### Needs

While these proposals are certain to encounter particular objections during the months ahead, they are broadly consonant with the needs of schools and of their students. There will, for example, be genuine disappointment, even anger, that the proposals for school science-teaching up to 16 suppose that physics, chemistry and biology will not be taught as if they were separate pieces of human wisdom. (The plain truth, of course, is that there are simply not the teachers to extend present specialist practices to all schools; almost certainly, there will not be enough qualified teachers to man the proposals now outlined for science and technology.) But two issues are neglected in the documents now published, of which one is that of the use to be made of the school-based attainment tests (at ages 7, 11 and 14) which are to accompany the new arrangements.

In the history of British education, attainment testing is understandably a contentious issue. Once upon a time (but the late 1950s were not so long ago) attainment testing at 11 was a means by which students were shunted from one educational stream to another, some offered the opportunity of winning a higher education six or seven years later and the rest condemned to a less demanding and stimulating curriculum in different kinds of schools. That system was not merely socially divisive but a waste of human resources (and a device indirectly of choking off demand for higher education). The unwinding of the system in British secondary schools during the 1960s and 1970s was an educational upheaval complicated by the determination of teachers to confuse the practice of comprehensive education with the supposition that all students are equal in interest and aptitude.

The big danger now (even for the introduction of the national



curriculum on which the government has properly set its heart) is that good sense will be undermined by the government's equal fondness for arrangements under which schools will be able to opt for independence from the local authorities that now control them, often no doubt to pursue selective educational policies of their own.

That will be contentious enough, but will be doubly so if the attainment tests now being designed are used not merely as determinants of students schooling (which is proper), but as ways of telling which schools they physically attend. That is why the government should be required to declare, in advance of the consultations on which its new array of curriculum councils is embarking, that it intends to make the national curriculum what the name implies. The curriculum, as modified in the light of changing circumstances, will realistically be a sign of what school-leavers have studied and have been helped to understand. If a student's attainment on some test is unsatisfactory, he or she will be helped to acquire the missing understanding. Attainment tests must therefore be educationally diagnostic, not instruments of educational apartheid. Without that undertaking, the national curriculum will be meaningless.

But how can such an undertaking be given when the resources available to the schools are as meagre as at present? Teachers complain at the lack of funds, but the more serious shortage is of people. Gifted mathematics teachers are too few to satisfy the needs of the modest curriculum proposed by last week's mathematics panel. Although the theme of the science panel is that science and technology should properly be a part of general education, it is nowhere demonstrated that instruction can be safely left in the hands of generalists. Yet that will be the result if, in the British educational system as a whole, the decline of interest in science courses in higher education continues at the alarming pace of the past few years. Despite recent contractions of capacity, British universities in particular are increasingly hard-pressed to fill the available science places with enthusiastic occupants. In short, there is a danger that the success of the best part of the Educational Reform Act will be vitiated by the separate provisions that have blighted higher education. □

## Prospect for US sciences

*Which two US parties' policies on science will best control the deficit?*

THERE was nothing in last week's convention of the Republican Party at New Orleans overtly intended to win votes at the election on 8 November, and very little at the earlier Democratic convention at Atlanta. Mr Michael Dukakis had a few uplifting things to say about the need to restore respect and self-respect to teachers as a whole and about the need to make full use of the inventiveness of laboratory people, but none of that can be counted as vote-catching. Left to themselves, both parties would probably follow the present administration in seeking to back the US research enterprise to the hilt, but with a few important differences (see p.639). What stands out from the conventions so far is that neither party has yet paid much public attention to the ways in which the post-election policies of the United States will be determined by considerations beyond the control of whatever party's champion is elected.

The unspoken questions so far are financial, and centre on the two still-huge deficits with which the United States is lumbered. The federal budget deficit (roughly \$150,000 million this year) is a measure of the degree to which the administration cannot pay its own bills from its revenue and is forced to meet the gap by borrowing instead; the external deficit, now fluctuating from month to month in the wake of the dollar devaluation but back to more than \$12,000 million last month, is a measure of the degree to which the borrowings are made from lenders overseas.

In principle, as the present administration has been fond of

saying, there is no reason why the two deficits should not persist. The precondition is merely that borrowers should be prepared to keep on lending until the US economy (and, with it, the yield of statutory taxes) grows to the point at which the federal deficit is once more in balance. The cost, of course, has been considerable. The United States is now the largest debtor nation (by a factor of three) while the devaluation of the dollar in the past nine months has made it easier for people overseas to buy US industrial enterprises at knock-down prices. But will foreigners keep lending to keep the US administration afloat? It has always been on the cards that their mood would change with the administration, somewhere between 8 November and mid-January next year; this likelihood has been increased by the increased interest rates now to be had elsewhere.

Even the past year's experience should make it plain that the changing economic pattern can seriously affect the federal government's capacity to support basic science. Once again, for example, the administration's laudable goal of doubling the budget of the National Science Foundation has been attenuated by the congressional agreement that the budget would be kept within limits derived from the old Gramm-Rudman rules (but less rigorous than those originally prescribed). Like it or not, science spending is discretionary spending, and must share with, for example, the military budget whatever cash is left over after statutory obligations (social security and the like) have been met. To judge from what the two presidential candidates have been saying, Dukakis will have the greater freedom to manoeuvre for the benefit of science: he believes that military spending could be reduced, while Mr George Bush has tied one hand behind his back by promising not to increase taxes. But it is too soon to know what will happen when one candidate or the other confronts reality.

That is why there is now an urgent need for some kind of contingency plan that will at once allow the federal deficit to be reduced to something like zero over a well-defined period (too sharp a reduction would be deflationary, and is in any case not required by lenders overseas) and which will allow some slack in the system for supporting basic science and its application, the need for which is crucial to the continuing renewal of the US economy. The difficulty, in the United States, is that some spending apparently peripheral to economic causes, including both military spending and the Strategic Defense Initiative, is far from irrelevant to the civilian economy. Companies commissioned one year to devise a customized chip for some esoteric spacecraft may, the next, find themselves incorporating the same device in some well-selling item aimed at consumers. But it is an assault on logic to suppose that practical benefits obtained as by-products of military projects are more economically won than those that derive from well-directed programmes of research. The presumption must be that the United States will gain more cheaply from the support of innovative research groups than from its extravagant ventures into future military and space technology.

This does not beg the strategic question, certain to dominate the next two months, of whether US military strength is a prerequisite of an accommodation with the Soviet Union on strategic arms control. It is merely that the change of an administration (not necessarily a change of party) provides an opportunity for disowning past commitments. Thus, it is even possible that Bush, if elected in November, could reasonably decide that there is no great need to hurry with the planned space stations and that the Strategic Defense Initiative would profit from being allowed to mature in the minds of its creators a little longer. Cutting loose from a few ventures of that kind for the sake of being able to spend a fraction of the funds (say a quarter) on basic science would be a good bargain, domestically and perhaps internationally as well. For whatever the election brings, the next administration will be the first in the United States that will have to learn that counting pennies is not merely a virtue but a necessity. □



# Presidential candidates find a little time for science

- Attitudes to SDI, AIDS, separate candidates
- Space exploration still catches voters

## Washington & Boston

WITH Vice-President George Bush's choice of Senator Dan Quayle as his running mate at the Republican Party Convention in New Orleans last week, the battle lines for the November US presidential election are finally drawn.

Scientists can be sure of one thing: scientific issues will not play a direct part in deciding whether George Bush or Massachusetts Governor Michael Dukakis will become the next president of the United States. But scientific interests do lie behind two of the three key issues — the Strategic Defense Initiative (SDI), AIDS and abortion — on which the candidates have very different views.

In the Democratic platform, science does not rate more than a passing mention. The Republicans, in contrast, give science and technology a full-page treatment and pledge support for all of the big projects — including the superconducting supercollider and increased funding for the National Science Foundation — linked with President Ronald Reagan's administration. But Bush has yet to make science the subject of a campaign speech.

Space exploration, the single scientific subject likely to capture the voters' imagination, has rated that honour. Bush is a supporter of the space station and, speaking at NASA's Space Flight Center in Huntsville, Alabama, earlier this month, he said he would consider backing a manned mission to Mars.

That did not stop Dukakis, speaking at NASA's Marshall Space Flight Center last Thursday from blaming the Republicans for letting the US space programme slide. Dukakis repeated his new-found support for the space station.

That support had emerged just as his running mate, Texas Senator Lloyd Bentsen, visited NASA's Houston Space Flight Center. With the massive Houston centre in his home state, Bentsen has been a long-time friend of NASA and has pushed for extra funds while a member of the Congressional committee that oversees space activities.

Dukakis has been equivocal about the space station. His hand was finally forced when, last week, Congress agreed to provide \$900 million for the space station in fiscal 1989. As half the funds could be vetoed by a new president, a group of congressmen pressed Dukakis to make his views clear.

Dukakis's attitude to the space station

now appears to differ from Bush's only in where he would find the funds to support it. Dukakis claims this can be paid for by cuts in the defence budget, particularly in funding for SDI.

On SDI, Bush and Dukakis find an issue on which they differ clearly. Bush has pledged to continue funds for the programme, but without Reagan's near-zealous support, many suspect that SDI may languish even in a Bush administration. Some indication of a new direction for SDI came from Senator Bob Dole at the Republican convention, when he spoke of the system as one that protected against an accidental missile launch, rather than as a defensive shield in a full-scale nuclear war. But Bush is running on the platform of continuing SDI.

Dukakis has repeatedly opposed SDI.



Presidential candidates — George Bush and Michael Dukakis.

In a speech on 14 June to the Atlanta Council in Washington, DC, Dukakis said "We don't need the Strategic Defense Initiative, we need the Conventional Defense Initiative". He also voiced opposition to several other pet Republican defence projects, including the rail-mobile MX missile.

The SDI programme is, however, so institutionalized that it may be hard to scrap. Research and development may continue, even if on a reduced scale. Dukakis's call for a Conventional Defense Initiative centres on his endorsement of a proposal by Joseph Nye, professor at the Kennedy School of Government at Harvard University, for a beefed-up conventional defence in Europe. Nye is a candidate for Secretary of Defense in a Dukakis administration.

Dukakis has also said that money taken from military research projects could be used to strengthen civilian research and that he will appoint a science adviser who will report directly to him. One person tipped for the job is Lewis Branscomb, former IBM chief scientist and now director of Science, Technology and Public

Policy Program at the Kennedy School of Government.

On the controversial topic of AIDS, both Democrat and Republican candidates claim they will spend whatever it costs to develop new drugs. Both party platforms specifically call for expedited FDA approval of new AIDS drugs. But on how to contain AIDS and treat its victims, big differences emerge.

While Bush has confessed difficulties in following scientific issues, Dukakis has emerged with an image as a successful 'technocrat', thanks to the highly touted 'Massachusetts Miracle' — a mainstay of Dukakis's campaign speeches. Massachusetts is second only to California in the number of high-technology companies it has attracted. Boston's route 128 — 'America's Technology Highway' — is a key computer business centre.

But critics of the Dukakis administration say he does not deserve much of the credit for the economic success of Massachusetts. One report issued almost a year ago charted the major role of military research and development monies in the state's success. And much credit must go to the many top universities and research institutes in the Boston area.

While Dukakis has the image of a technocrat, it is Bush who has offered something concrete to small high-technology business. He has been vocal in support of the permanent extension of the research and development tax credits scheduled for elimination in 1988. The Democrats remain silent on the issue.

On environmental issues, the Democrats have the better image. Bush has recently begun to try to change this, talking more forcefully about ocean dumping and acid rain. He also switched policy on offshore oil drilling, an issue on which he had angered Californians.

For Dukakis, environmental issues have always played a strong role in his rhetoric, if not in his actual policies. Dukakis is now calling for initiatives that go far beyond those of Bush in cleaning up the environment. But while he has made strong statements about ocean dumping and protection of the coastline, Boston Harbor, the most polluted in the nation, remains a large thorn in his side.

Dukakis has taken a tough-sounding stance on acid rain and has called for strict national standards to cut emissions of sulphur dioxide and nitrogen oxide. In Massachusetts, however, environmentalists claim that Dukakis has done little to support new state smog-control plans.

Dukakis's backing for environmental issues has also led him to oppose strongly the Seabrook nuclear power plant. He has been a critic of the nuclear power industry in general, but has not publicly opposed nuclear power outright. In contrast, Bush has been a consistent supporter of nuclear power. **Alun Anderson & Seth Shulman**



## NSF to get big budget increase but not as much as it wants

### Washington

THE US National Science Foundation (NSF) is set to receive a 9.8 per cent increase in its budget following agreement between House of Representatives and Senate appropriations committees. The appropriations bill, which gives NSF a total budget of \$1,885 million for fiscal year 1989, has now gone to the White House for signing.

Although the rise is only half that originally requested by the Reagan administration, it is above the current rate of inflation and should allow NSF some room for expansion. A big increase was earmarked for efforts to improve science and engineering education (22.8 per cent). But the bill refused NSF's request for a separate \$150 million budget, spread over five years, to fund the proposed new science and technology centres. The centres are intended to foster multi-disciplinary collaborative research and help move the results of basic research from laboratory to industry. Hopes that

the first centres could be established this fiscal year have already been disappointed.

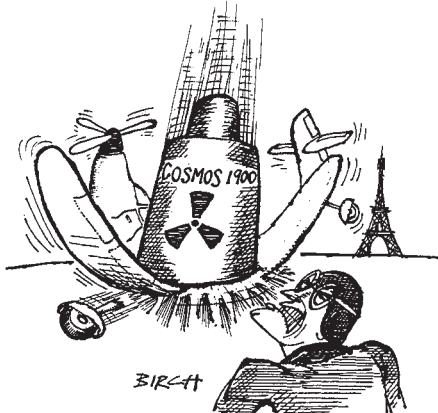
The appropriations bill does provide for "a limited number" of new centres to start up next year with funds from NSF's core \$1,583 million appropriation for research and related activities.

That account provides support for individual researchers as well as for already-established science and engineering centres, supercomputer centres and many other programmes. NSF is expected to decide soon how much it can spare towards the new centres and how many of the fifty proposals it has already received can be supported.

An upper limit is eventually likely to be set in the authorization bills which direct how appropriated funds may be spent. Both House and Senate authorization bills failed to complete passage before the summer recess. At present, the Senate version authorizes expenditure on science and technology centres in fiscal year 1989 of \$30 million, and the House version authorizes \$20 million. Compromise will have to be found on this issue and that of whether a new ice-breaker ship for the Antarctic Program should be built in the United States or ordered from abroad at a lower cost (see *Nature* 330, 789; 1988). The appropriations bill suggests that the ice-breaker should be built in the United States unless the cost is more than 50 per cent greater than that for a foreign ship.

Alun Anderson

## French keeping their heads down



### Paris

WHILE 300 gendarmes scan the night sky to try to pinpoint a mysterious small aircraft that has been skimming Paris rooftops for the past week, a number of government departments have been preparing contingency plans in case the wayward nuclear-powered satellite, COSMOS 1900, should crash on French soil. Although the risk is considered to be extremely low (about one in one thousand), joint emergency plans have been drawn up by the Civil Defence corps, the Atomic Energy Commission (CEA) and the Central Service for Protection against Ionizing Radiations (SCPRI). These include the mobilization of 35 helicopters, and telling the public to keep away from suspicious metal objects.

Peter Coles

## Liberal stance on embryo experiments

### Sydney

NEW South Wales has been urged to adopt less stringent regulations on *in vitro* fertilization research than the controversial regulations agreed in the neighbouring state of Victoria.

In a report presented by the New South Wales Law Reform Commission to the state parliament, the commission urges that all types of research on human embryos be permitted until 14 days after fertilization, the usual time of implantation. In Victoria, research is limited to a 22-hour period after fertilization.

Important reasons behind the committee's liberal stance were acceptance of the proposition that embryos need not necessarily be given the full rights accorded to people, and the realization that the poor success rate of *in vitro* fertilization techniques can be improved only through further experiments.

Charles Morgan

## Indian earthquake

### Washington

LAST Sunday's earthquake in the Indian state of Bihar, near the Nepalese border, had a magnitude of at least 6.5 on the Richter scale and, according to early reports, has killed as many as a thousand people. At least 100,000 homes have been destroyed by the earthquake. Reports from rural regions have not been received yet and are bound to push the death toll even higher.

The earthquake was registered on seismographs at US Geological Survey stations and was officially estimated at 6.7 on the Richter scale. Estimates from New Delhi were slightly lower.

The earthquake occurred on the same fault line, in the Himalayan foothills, as the magnitude 8.3 earthquake that killed 7,500 people in 1934. In most parts of the world, an earthquake of magnitude 8.0 or more would be expected to relieve most of the geological stress in the area, and thus lead to a fairly quiet period lasting perhaps for centuries. But at the Himalayas, where the Indian subcontinent is running into Asia, the relative motion of the continental plates is extremely rapid — close to two inches a year. According to Bruce Bolt of the University of California, the 50-year interval between the 1934 earthquake and last week's cannot be called unusually small.

The death toll from Himalayan earthquakes is much higher than would occur in California, for example, where similar shocks have been registered. The San Francisco earthquake of 1906 was of magnitude 8.3, and killed about 700 people; a 1971 earthquake in the San Fernando valley, with magnitude 6.7, killed 63. The difference is partly explained by the higher density of population in India and partly by the topography: Himalayan earthquakes frequently trigger large landslides.

The Indian government has been making attempts in recent years to educate the population about earthquakes and to implement new safety codes for buildings, but poverty and the remoteness of the population centres make earthquake awareness programmes of California sophistication impossible. An earthquake of magnitude 8 or greater occurring in the Bihar-Nepal region today might, according to some experts, cause up to a million deaths.

With the help of the United States, India has begun to install a network of strong-motion accelerometers along its northern borders. A number of high dams have been built in the Himalayan foothills over the past few years, and the Indian government is concerned to protect its investment by studying the pattern of seismicity in the area, as well as to provide earthquake warnings for lower inhabitants.

David Lindley

# US-Soviet joint monitoring of nuclear explosion in Nevada

*Washington & Las Vegas*

A HISTORIC step towards a nuclear test-ban treaty was taken on 17 August when thirteen US and seven Soviet scientists jointly monitored the explosion of a 150-kiloton US nuclear device at the top-secret US Nevada Nuclear Test site.

The event was designed to give scientists from both nations the chance to match their methods of monitoring the size of nuclear explosions. Next month, US scientists will attend a similar test at the Soviet Semipalatinsk nuclear site.

If agreement can be reached on how to measure the size of nuclear explosions, a major impediment to ratification of two treaties limiting the size of nuclear tests will be removed. These are the Test Ban Treaty of 1974 which limited underground explosions to 150 kilotons, and the Peaceful Nuclear Experiments Treaty of 1976.

According to Geneva nuclear test-ban negotiations C. Paul Robinson of the United States and Igor Palenikh of the Soviet Union, who were present for the Nevada test, the joint monitoring programme may eventually lead to a comprehensive test-ban treaty.

In the past, the US administration has accused the Soviet Union of breaking the 150-kiloton unratified test-ban limit. The Soviet Union has denied these charges, saying that the seismic monitoring methods it favours show the limit has not been violated. But the US side has argued that only the CORTEX method it favours will accurately measure the size of nuclear blasts. CORTEX is very intrusive compared to seismic methods as it involves placing a long cable in a borehole just alongside the shaft in which the nuclear device is detonated and measuring the rate at which the cable is crushed (see *Nature* 326, 818; 1988).

## Boost for shuttle

*Washington*

THE US space shuttle's rocket booster was successfully test-fired on 18 August at Morton Thiokol Inc.'s test facility, clearing the way for the shuttle *Discovery* to be launched in September or October. In order to test rigorously the redesigned booster's safety features, 14 flaws were deliberately introduced into the booster rocket before the test. These included damage to the O-rings that seal the joints between booster segments. But even with unrealistic levels of damage hot gas did not appear to escape from the joint seals. Accidental escape of hot gas triggered the explosion that destroyed the space shuttle *Challenger*.

Alun Anderson

The use of CORTEX inevitably means that Soviet and US scientists have to visit each others' test sites. Despite earlier objections to such visits, the Soviet government last year changed its policy and proposed a bomb test exchange in which both sides would use essentially the same CORTEX techniques.

In the Nevada test, identical cables from the monitoring borehole led to separate US and Soviet instruments. Shortly after the blast, Viktor N. Mikhailov, Soviet team leader at the Nevada Test Site, announced that Soviet instruments had worked and data had been recorded. The room burst into applause.

Raw data gathered on 17 August will be analysed along with information collected at next month's Semipalatinsk test. Only then will it be clear if Soviet and US instruments are yielding the same results. If the tests prove successful, they are expected to be followed by negotiations over how often each side should allow tests to be monitored. Agreement should make it possible for the Threshold Test Ban Treaty to be ratified.

At the same time as the on-site monitoring was carried out under the official US-Soviet agreement, groups of Soviet and US scientists sponsored by the US National Resources Defense Council (NRDC) and the Soviet Academy of Sciences made unofficial off-site measurements using the seismic techniques that

have been favoured by the Soviet Union. NRDC has been pressing the US administration to accept seismic monitoring and has sponsored visits of US experts to the Soviet Union (see *Nature* 328, 192; 1987). NRDC experts hope that their new results will persuade the US administration that seismic methods yield the same results as on-site monitoring.

Alun Anderson & Mary Manning

## Scientists in trouble

*Las Vegas*

THREE unidentified US geologists will not participate in the Soviet Union's joint verification experiment scheduled for 14 September at Semipalatinsk after Soviet inspectors discovered unauthorized site samples in their luggage.

The Soviets filed a protest on 25 July, but did not proceed with a formal complaint after officials agreed to remove the trio from the experiment. Department of Energy (DoE) officials called the rock, soil and wire samples "souvenirs". DoE sent a team of experts to the Soviet Union to prepare for an unprecedented exchange of data as the United States did for separate nuclear blasts measured by both super-powers.

The US scientists had soil, rocks, barbed wire and a hammer in their luggage. Soviet authorities said the materials violated the terms of an agreement signed by super-power leaders at the Moscow summit.

DoE Nevada Operations Manager Nick Aquilina said after the successful US test on 17 August that the collectors did "something foolish".

Mary Manning

## Landsat keeps an eye on continuing US drought



The Mississippi River, three miles north of Greenville during an average year (left, 3 June 1985), when river gauge readings measured 26.8 feet, and on 11 June this year, when gauge readings were at a record low of 12.1 feet. At point A in the images the river has decreased in width by a quarter of a mile, and the water level in the oxbow lake to the right is drastically reduced. In the main water channel B, strings of barges can be seen in the 1988 picture, waiting for an increase in water level. At C an increase in size of the large sandbar has closed off some of the river channels just east of Arkansas City. The bright green line D is the system of control dykes.



## Future of satellite technology being rethought everywhere?

### Delhi & Paris

THE partial power failure which crippled India's \$140 million multipurpose satellite, INSAT-IC, within a week of its launch by a European Ariane rocket last month has come as a blow to the Indian Space Research Organisation (ISRO). But it also highlights a growing uncertainty in the international commercial satellite market which is encouraging both manufacturers and users to look again at Earth-based telecommunications systems.

Although the cost of replacing INSAT-IC should be borne by its manufacturer, Ford Aerospace Corporation, its partial failure has upset India's plans to expand its national television and radio, telecommunications and meteorological networks. Fortunately, six of the satellite's twelve transponders are working perfectly, allowing a vital meteorological early-warning system to come on stream and providing back-up for INSAT-IB, which is due to be retired in mid-1989.

The future of the \$350-million INSAT system now depends solely on INSAT-ID, the last in a series of imported satellites, to be launched by a US McDonnell-Douglas Delta rocket next year, it is hoped before INSAT-IB expires. Satellites of the INSAT-2 series, which will be bigger and heavier, will be built by ISRO and are scheduled for launch by Ariane rockets in 1990 and 1991.

Unable to expand the INSAT system as planned, an Indian interministerial committee is currently reworking communications priorities in order to make best use

of INSAT-IB and the crippled INSAT-IC. But this latest failure will not only add to concern over mounting costs of India's indigenous space programme. It also adds to an increasing nervousness worldwide regarding the costs and risks of satellite technology.

Now that so many reliable civil launch vehicles are once again available — with Chinese, Japanese and Soviet launchers challenging US and European domination — it is the satellite industry itself that is being threatened. Paradoxically, this is due, on the one hand, to the long life of successful satellites and, on the other, to an increasing number of failures. India's plans to develop its own rocket launcher, already dogged by two failures in a row, may now turn out not to be viable and Prime Minister Rajiv Gandhi has ordered an in-depth review of ISRO's launch vehicle programme.

It is unlikely that India's domestic space programme will justify the cost of developing an indigenous launcher, and a stagnating world satellite market promises slim pickings even for those who already have a share. In the first half of 1987, only two satellites were ordered from major US constructors — leaders in the field — compared with twelve in 1985. Current trends suggest a levelling-off of demand for civil satellites to about 10–15 per year in the early 1990s. And if some sectors, such as direct broadcasting and telecommunications, are still expanding in Europe and Japan, improvements in long-distance fibre-optics cables are making risky satel-

lite technology appear increasingly unattractive. Private sector investment in fibre-optics technology has already overtaken investment in satellite technology, even among leading satellite constructors.

Even the outwardly attractive scientific and technological advantages of a manned space platform, such as microgravity conditions, are coming under scrutiny because of the setbacks caused by the withdrawal of the US space shuttle. With a backlog of around 140 experiments planned for Spacelab, US companies are wondering whether microgravity projects will be commercially viable, and McDonnell-Douglas has already started to develop Earth-based laboratories for its revolutionary electrophoresis process.

K.S. Jayaraman & Peter Coles

## South Africa states NPT conditions

### Oxford

THE South African government has declared that it would find it difficult to be a signatory to the Nuclear Non-proliferation Treaty (NPT), unless it has the right to buy and sell uranium and exchange nuclear technology internationally. In a joint statement released in Cape Town last week on their return from Vienna, Minister of Foreign Affairs Pik Botha and Minister of Technology Danie Steyn claim that these rights are implicit in the treaty. In Vienna, South Africa was urged by the Soviet Union, Britain and the United States, the main guarantors of the treaty, to accede to the NPT, thereby opening all its nuclear installations to inspection by the International Atomic Energy Agency (IAEA).

The ministers' precondition is clearly aimed chiefly at the United States, which imposed an embargo on the importation of South African uranium in January 1987, although other NPT signatories which restrict the importation of South African uranium include Norway, the Soviet Union, Canada, Australia and New Zealand. South Africa is particularly worried about its declining trade with the United States; last year, South African exports to the United States were worth just over half their average annual value between 1983 and 1985. Even since the law was tightened up in the middle of last year, the United States still permits the importation of uranium mined in South Africa, but enriched by a third party. Similarly, the Soviets are also said to countenance the import of uranium of South African origin.

A revocation of the US ban appears extremely unlikely in the current political climate. Thus South Africa's action is likely to be seen as a red herring, particularly as it refused to sign the NPT for almost 20 years before the implementation of the ban.

Michael Cherry

## Court rules for demolition of West German nuclear plant

### Munich

A WEST German court decision has cleared the way for the demolition of a large nuclear power plant. The decision on 8 August by the Administrative Court at Regensburg rejected a citizen's claim that radiation released during demolition could endanger the local population.

Researchers at the Nuclear Research Center at Karlsruhe (KFK) are in the third year of the demolition of the nuclear power plant, owned by KFK, at Niederachbach, near Munich. The reactor is a heavy-water-moderated, carbon dioxide-cooled pressure tube reactor with 100 MW generating capacity.

KFK will not complete demolition until 1992 at the earliest. Remote-control equipment powerful enough to move several tonnes of material will be used for the first time. Demolition will cost DM140

million. The reactor cost DM230 million and was built between 1966 and 1972.

Niederaichbach citizen Franz Kohout began his legal challenge of the demolition in 1986. The court could have halted demolition of the reactor core, but it ruled that the amount of radiation released would not exceed the allowable level of 30 millirem per year.

Beginning in 1990, a KFK team will dismantle up to 500 tonnes of radioactive steel and 500 tonnes of concrete from the core and shielding of the plant and pack it into shielded barrels for transport to permanent storage sites. Roughly 1,700 tonnes of waste steel that is only slightly radioactive will be recycled in Karlsruhe and used in nuclear shielding or storage containers. Fuel elements have already been removed from the plant.

Steven Dickman



# US report accuses NSF of polluting polar environment

## Washington

US ANTARCTIC research bases run by the National Science Foundation (NSF) are polluting the unspoiled polar environment in defiance of national and international regulations, according to a report released last week by the Washington-based Environmental Defense Fund. But NSF officials say that "increasing and extensive efforts to improve performance" have been under way since the late 1970s and the situation is nothing like as bad as the report claims.

Bruce Manheim, the report's author, says that McMurdo Sound where the largest US base in Antarctica is located is more heavily polluted than virtually any part of the United States' own coastal waters. Polychlorinated biphenyls (PCBs) and heavy metals — including lead and cadmium — are now found in the tissues of Antarctic penguins and seals.

Manheim describes it as "ironic" that the pristine conditions that make much Antarctic research possible are being harmed by the day-to-day running of research bases. According to the report, raw sewage is discharged into the sea, combustible wastes are burnt without emission controls, non-combustible wastes are often dumped into the sea and the air is polluted by diesel-powered generators.

"Other nations are doing a better job than the United States", says Manheim. Australia removes all wastes from the Antarctic; Poland uses a filtered sewage system and both New Zealand and Australia control incinerator emissions. Britain's practices are said to be the same as those of the United States.

Jack Talmadge, director of polar coordination and information at NSF, describes the claims of large-scale pollution in McMurdo Sound as "preposterous".

Some of the criticisms in the report are out of date, he says. Others "may be legitimate" but do not mean that the United States does not care about the Antarctic environment. NSF has taken informal advice from Australia which Talmadge describes as the "leader" in minimizing Antarctic pollution and is looking for funds to help clean up the McMurdo base.

The report also claims that NSF has misled the public and Congress. In 1978, Congress enacted the Antarctic Conservation Act which gave NSF responsibility for promulgating regulation to prevent pollution, but NSF has never gone beyond issuing guidelines. The act also directed NSF to protect all Antarctic marine life. But the report says that activities that may affect marine life are not closely regulated.

The report points out that the number

of bases in Antarctica is likely to grow as more nations seek to "conduct substantial scientific research activity" so that they become eligible to accede to the Antarctic Treaty as consultative partners. That status is necessary for a say in what happens to Antarctica and its resources. The first steps to open Antarctica to mining

were taken at this year's treaty meeting (see *Nature* 333, 588; 1988).

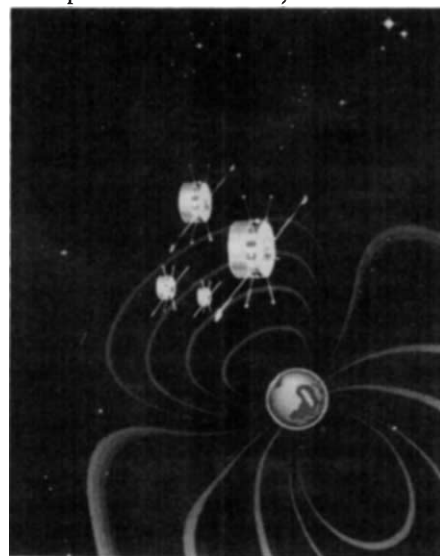
Exploitation of Antarctica's mineral resources is still a long way off, however. A more pressing problem is the growth of tourism. Talmadge says that next season he expects 3,200 ship-borne tourists to visit the Antarctic Peninsula, where most bases are located. The total number of scientists there is probably less than 500. And, unlike scientific activities, tourism is almost totally unregulated.

Alun Anderson

# UK space community cheered by contribution to solar research

## London

BRITAIN'S space science community is cheered this week by the size of the country's contribution to the SOHO and CLUSTER missions. This eight-year programme is the first major step in the European Space Agency's (ESA) 20-year programme Horizon 2000. Britain is to spend £56 million on the satellites and £20 million on instruments. Announcing Britain's contribution, Arthur Pryor, head of the British National Space Centre said it was "an exciting and challenging opportunity for the whole science community." This solar-terrestrial science project will cost about £750 million; just over half the cost is being met by ESA and the rest by NASA (National Aeronautics and Space Administration).



Artist's impression of CLUSTER (Photo: BNSC).

The aim of the SOHO mission is to understand the nature of the Sun's atmosphere and how it is heated, to study how the Sun expels material into space and to probe the Sun's interior by studying its vibration patterns. CLUSTER, which will consist of four identical spacecraft flying in close formation in Earth orbit, will study the Earth's magnetosphere and the

solar wind. Both missions will be launched in 1995 and should be able to provide useful data four months after launch. CLUSTER will be complemented by two NASA spacecraft (WIND and POLAR) and a joint US-Japan satellite (GEO-TAIL), which will explore other parts of the magnetosphere.

Britain surprised its partners in Europe when it agreed to provide the £20 million to support the British instruments which were selected for the missions. In the European space community, Britain made itself unpopular by its refusal to consent to a five per cent annual increase in the ESA budget beyond 1989. But a condition of the country's involvement in the Columbus space station programme was that it would reconsider that veto, and it has until the end of the year to do so.

Christine McGourty

# New data centre

## Washington

THE US National Science Foundation (NSF) agreed on 18 August to provide \$5.5 million over five years for a new National Center for Geographic Information and Analysis. Its aim will be to develop more powerful methods of dealing with the explosion of data from such sources as remote-sensing satellites. It is to be run by a consortium of the University of California at Santa Barbara, the University of Maine at Orono and the State University of New York at Buffalo.

Alun Anderson

# European genome plan

## London

THE European Commission is undertaking a £20 million research project to sequence the human genome. The commission will provide half of the funds; the rest will be contributed at a national level. The project, 'Predictive Medicine', must still be approved by the Council of Ministers and the European Parliament. The 3-year project starts in 1989.

C. McGourty

# Hungarian establishment now opposes hydroelectric project

London

THE Hungarian Academy of Sciences has asked the government for permission to publish its three reports on the probable environmental effect of the Gabčíkovo–Nagymaros hydroelectric project. This is the latest in a sequence of events that, since the conference of the Hungarian Socialist Workers' Party two months ago, has transferred open opposition to the scheme from fringe environmental organizations to establishment bodies.

The project, which involves the diversion of the Danube through Slovakia, north from its present channel, and the construction of two major 'peak-hour' hydroelectric plants at Gabčíkovo in Slovakia and Nagymaros in Hungary, was originally intended as a joint Czechoslovak–Hungarian undertaking. But three years ago when the Hungarian side deferred its part due to lack of money, the Austrians, who had been blocked by their own 'greens' from building a similar peak-hour plant at Hainberg, offered to finance the Hungarian part of the structure.

By this time, there was a growing groundswell of opinion against the dam in Hungary (the first two highly critical academy reports had already been drawn up). The then Hungarian leader, Mr Janos Kadar, preferred to use quiet diplomacy with the Czechoslovak side, and, when the

latter proved unwilling to cancel, to plead poverty. The Austrian involvement, in effect, called his bluff, and construction began on the Hungarian side too. All that the academy could do at that stage was to negotiate additional funding to offset the environmental threat.

The party conference of 22–23 May 1988 brought a change of leadership. The day after Kadar was replaced by the incumbent prime minister, Mr Karoly Grosz, the recently formed 'official' Hungarian Society for the Protection of the Environment issued a call for a new appraisal of the economic and environmental costs of the scheme. Four days later, there was a demonstration in Budapest against Austrian involvement. Next, a member of parliament, Mr Zoltan Kiraly, called for a full parliamentary debate into the dam. This will take place during the autumn, and the preparatory briefing of MPs has now begun, including, significantly, figures on the cost of cancelling the whole project, or of completing the power station at Nagymaros without the projected dam.

In mid-July, the president of the Hungarian Academy of Sciences, Academician Ivan T. Berend, met Dr Bela Liptak, head of the Foundation for the Protection of the Hungarian Environment, an association of expatriate Hungarian scientists based

in the United States. Together they wrote a letter to the Hungarian government, setting out the case against the dam. This letter has since been heavily attacked by the Slovak daily *Pravda*, which claimed that Liptak's figure contradicted those of the Hungarian and Czechoslovak specialists. The figures, however, come from the unpublished academy reports, while according to Kiraly, members of the Slovak Academy are becoming increasingly worried that the Nagymaros dam will drown or swamp much of Slovakia's best agricultural land.

Vera Rich

## Japan subsidizes superconductors

Tokyo

JAPAN'S Ministry of International Trade and Industry (MITI) is planning to double government subsidies for the International Superconductivity and Technology Center, established in January with massive funding from the private sector (*Nature* 331, 105; 1988). But, despite an enviable budget, there are still few foreign companies willing to join the centre.

Membership of the centre for companies comes in two forms: 'general' members can gain access to an information centre and symposia for a down-payment of ¥2 million (\$15,000) and an annual subscription of the same amount, while 'special' members must fork out an additional ¥100 million (\$750,000) plus ¥12 million a year to join the centre's research laboratory which is due to open in Tokyo in October.

More than a hundred companies have joined the information centre and 45 have signed up for the laboratory, bringing in more than ¥5,200 million (\$39 million) in the first year of operation. In addition, MITI has provided the centre with about ¥500 million in subsidies in this fiscal year and the ministry plans to double this to about ¥1,000 million in fiscal year 1989, according to Professor Shoji Tajaka who will head the research laboratory. MITI will submit a budget request for the subsidies to the Ministry of Finance at the end of this month. But only five foreign companies, four from the United States and one from Britain, have joined the information centre and none has joined the research laboratory.

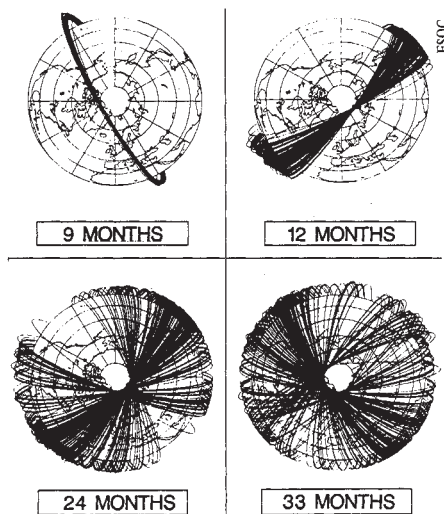
Apart from cost, Tanaka thinks that problems over patents are deterring participation of foreign companies. The 45 Japanese companies that have joined have spent months hammering out an agreement on how to deal with this issue. Tanaka expects the agreement to be drafted by next month and it will then be translated into English and provided to potential foreign participants.

David Swinbanks

## Dangers of more debris in space

Paris

THE European Space Agency (ESA), in a report to be published later this year, has been taking a look at the problems of space debris which have recently been worrying the US Congress (*Nature* 334, 188; 1988).



Computer simulation of the evolution of fragment orbits from an explosion.

On the basis of computer simulation, ESA's working group on space debris estimates that there are between 30,000 and 70,000 small fragments orbiting the Earth, in addition to the 7,000 larger objects currently being tracked by the United States and the Soviet Union.

The working group was set up in 1986, under the chairmanship of Walter Flury, to look at collision hazards in the geostationary orbit, 36,000 km above the Earth, and the risks to spacecraft and manned missions in lower orbits. Of particular concern to ESA are the Hubble Space Telescope and the Hipparcos astronomical satellite, which are likely to suffer interference from light refracted off small particles, as well as the risks of damage through collision.

According to ESA's estimates, the Hubble Space Telescope will run a one per cent risk of colliding with a "large piece of junk", over 10 cm in diameter, during its 17-year lifetime. But collision with smaller debris, for which no figures are given, would also be serious. A pea-sized fragment orbiting at 5 km per second "could shatter a \$100 million satellite", said an ESA spokeswoman.

Peter Coles



# Growing reaction to ozone hole in Soviet Union

## London

*Sovetskaya Rossiya*, the newspaper that two years ago led the campaign to halt the southward division of the Siberian rivers, has now turned its attention to ozone holes. In an appeal to the USSR State Committee for the Protection of Nature, it called for a "ruthless struggle against those who violate ecological laws". A recent computerized study, it reported, had observed more than 50 cases of reduction in the concentration of ozone above major cities in the European part of the Soviet Union, including Moscow.

*Sovetskaya Rossiya's* concern differs considerably from the attitude to ozone holes over the past few months of the official Soviet news agency TASS, which has on several occasions suggested that ozone holes are a puzzling, but not anthropogenic phenomenon. In April, reporting that a team at the Soviet Molodezhnaya station in Enderby Land would be spending the forthcoming southern winter in carrying out the first-ever "comprehensive" ozone monitoring programme, it cited one of the team's leaders, Dr Aleksandra Chetverikov of the Volkov geophysical observatory in Leningrad, as saying that data gathered at Molodezhnaya during the summer of 1987-88 may refute the "ozone hole" theory.

A few weeks later, TASS produced a Georgian scientist, Dr Teimuraz Toroshelidze, who declared that ozone holes hold no threat of disastrous consequences for mankind; they are a normal natural phenomenon, resulting from turbulent flow round a high obstacle such as a mountain or the Antarctic plateau. Fluctuations of ozone concentration are frequent over the Caucasus, he said, and are caused by strong winds above 5 km over the northwestern part of the range. Above the Abastumani astrophysical observatory, where Toroshelidze works, an increase of wind velocity by 10 m per second reduces stratospheric ozone content by some 5 per cent, he said.

Even those scientists who have stressed the ultraviolet hazard of ozone holes, such as Dr Nikolai Elanskii of the Institute of Atmospheric Physics of the Soviet Academy of Sciences, seem to take an ambivalent stance in TASS interviews. Although Elanskii noted that the Soviet Union had signed the 1985 Vienna Convention on the Ozone Layer and the 1987 Montreal protocol on ozone-destroying substances, and by 1993 will have switched from freons to ozone-friendly hydrocarbons for domestic aerosols, he emphasized that the beginning of "intensive ozone destruction" coincided with the 1982 eruption of El Chichon, implying

that aerosols from the Mexican volcano were chiefly to blame.

Soviet scientists are eager to cooperate with their western colleagues in studying the phenomenon. According to Dr Igor Karol, head of the laboratory of climate theory at the Main Geophysical Observatory of the USSR in Leningrad, recent Soviet proposals for cooperation include a chain of monitoring stations (one in the Canadian arctic archipelago, one on Cape Barrow, one on Wrangel Island), simultaneous Soviet, Canadian and American trans-Arctic flights, and — most interesting — the 'internationalization' of one of

the Soviet drifting icefloe Arctic stations to include Canadian and US teams in 1989 or 1990.

Vera Rich

■ The results of the US expedition during the 1987-88 Antarctic summer, in which ozone concentration and atmospheric chemistry were mapped over the whole region, convinced even the manufacturers of chlorofluorocarbons that their products were implicated in polar ozone destruction. No clear indications of ozone loss over cities in the United States are apparent from satellite measurements, but according to Mark Schoeberl, of the NASA/Goddard Space Flight Center, local ozone concentrations fluctuate daily as weather conditions change, and a build-up of sulphur dioxide from urban pollution can look to the satellite like a stratospheric ozone loss.

David Lindley

# Arizona seeks to sidestep the endangered species law

## Tucson

CONGRESS, not regional wildlife biologists, will decide the best way to manage the habitat of an endangered squirrel if the University of Arizona and its congressional delegation have their way.

The 200 or so extant Mount Graham red squirrels, *Tamiasciurus hudsonicus grahamensis*, live on Mount Graham, an 11,000-foot-high peak in southern Arizona that is the proposed site for a University of Arizona observatory. Astronomers want to put seven telescopes, including the Columbus Project and the Max Planck submillimetre telescope, on the mountain and have spent more than a million dollars evaluating, planning and lobbying for the facility.

Frustrated at waiting for federal agencies to complete their assessment of the impact of land use on the squirrel population, and concerned that telescopes would be located elsewhere if construction were delayed any longer, the university asked its congressmen to designate the mountaintop an astrophysical reserve exempt from the assessment process.

Attempts to attach a rider to federal legislation before the recess on 12 August failed, but US Fish and Wildlife and Forest Service administrators in Washington are working with Arizona congressmen on the proposed legislation.

In July, the US Fish and Wildlife Service rejected the proposal to build seven telescopes on Mt Graham's twin peaks, but offered three alternative plans to the university and the National Forest Service, which has the final say on land use (see *Nature* 334, 188; 1988). The university accepted an option to build three telescopes on Emerald Peak alone, but wanted additional land reserved for four more telescopes, anticipating that the

objections could eventually be overcome. The university also objected to a demand that an old access road, through 'squirrel country', should be closed before a new one was built, as that would have delayed the start of construction until next year.

Complaining that it had already been waiting for four and a half years, the university asked the Forest Service to make a final decision by the end of August. Laurel L. Wilkening, vice-president of research at the university, told the Forest Service that one telescope had already been lost, and another (the Max Planck telescope) will soon be lost. "We cannot pursue any additional processes that would require further delay", she said.

But the Forest Service said a final decision could not be made for several months. Consultation with the US Fish and Wildlife Service would be necessary as well as public discussion and a new environmental assessment of the alternative Emerald Peak plan.

The university does not want to wait and has tried to enlist congressional help. Arizona senators John McCain (Republican) and Dennis DeConcini (Democrat) tried on 4 August to attach a rider to the Senate dire emergency supplemental appropriations bill which would have allowed immediate construction of three telescopes and land for seven. The four other telescopes were to be built after consultation with the US Fish and Wildlife Service. Although the language of the rider was never finalized, Randall A. Smith, a biologist with the National Forest Service said that the legislative actions "are taking the final decision away from the Forest Service and they are taking away from the public the right to comment".

Elizabeth Pennisi

## Jumping the greenhouse gun?

SIR—I do not know from what source John Maddox is quoting me in your issue of 7 July<sup>1</sup>, nor do I admit to jumping any gun. He believes I should know better. I believe I *do* know better.

Assuming that he is referring to my address to the World Conference on the Changing Atmosphere on 27 June 1988, in Toronto, then I *asserted* nothing, but argued as follows:

And can we detect the greenhouse effect in existing climatic records? (No approach) has yet given an unequivocal answer. But painstaking examination of the global surface and sea temperature records over the past century has begun to yield good results. Analysis by Jones *et al.* (1986), for example, shows a global rise of temperature between 1860 and 1980 of approximately 0.5 to 0.6°C (ref. 2). Villach 1985 (a WMO/UNEP/ICSU assessment) estimates that the greenhouse effect over the past century should have raised temperatures between 0.3 and 0.7°C (ref. 3). The observed signal is thus compatible with the predicted greenhouse warming . . . . Firmer answers to this question will have to wait a few years. But I should like to express the personal view, based on long experience rather than personal research, that we are indeed witnessing the beginnings of the process, and that the delegates did not come to the conference to chase a will-o'-the-wisp<sup>4</sup>.

I added, in a verbal aside, that I agreed with remarks attributed the previous week to J.E. Hansen, before a US congressional committee, that "it is time to stop waffling so much and say that the greenhouse effect is here"<sup>5</sup>. Of course I did not testify myself.

Until a short while ago, my own position (as chairman of the WMO/UNEP/ICSU Advisory Group on the Greenhouse Cases, *inter alia*) was much like Maddox's: that the evidence was too equivocal to do more than give a yellow alert to governments (for whom, incidentally, the Toronto Conference was run). I was an announced wait-and-see conservative.

It was the paper by Jones *et al.*<sup>2</sup> published in *Nature* that began to sway me towards the above position. This was reinforced by an update, also in *Nature*, establishing that 1987 was the warmest year on record, and that the 1980s clearly show a resumed upward trend<sup>6</sup>. Overall, there is an unmistakable quasilinear upward trend over the entire period 1860–1987, on which large interannual and interdecadal fluctuations are superposed, as Maddox indicates. This is true in both hemispheres.

Obviously this is no proof that the greenhouse effect is at work, nor will there be

any such proof that satisfies everyone, hence my words "personal view" not committing either of the two bodies I chair. But the trend fits what theory predicts, in the shape of several major general circulation models that have tested the hypothesis, and come up with the figures I quoted. All the difficulties and uncertainties cited by Maddox are, of course, well-known to those of us who play this game.

As a scientist, I will hence confine myself to saying that the best available explanation for the upward trend of surface temperature is the build-up of the greenhouse gases, and the response of the atmosphere to the change in optical behaviour. As an adviser to my government I have to be more explicit. I can and do tell them that they should base their environmental planning on the assumption that the greenhouse warming will continue and accelerate. There will always be conservatives who decline to go this far. At the age of 69 I can no longer afford to be conservative.

F. KENNETH HARE  
(Chairman)

*Climatic Planning Board of Canada,  
Toronto, Ontario,  
Canada*

1. *Nature* 334, 9 (1988)
2. Jones, P.D. *et al.* *Nature* 322, 430–434 (1986).
3. Villach 1985, *Report . . . on the Assessment of Carbon Dioxide . . . in Climatic Variations and Associated Impacts*, WMO-961, 75pp. (1986).
4. Hare, F.K. in *Report of the World Conference on the Changing Atmosphere*, Toronto, 27–30 June 1988 (in the press).
5. Highfield, R. *Observer*, 28 June 1988.
6. Jones, P.D. *et al.* *Nature* 332, 790 (1988).

SIR—In my book *Pesticides and Pollution*, published in 1967, I said: "So far most scientists have thought that CO<sub>2</sub> pollution was of little importance; it now seems possible that it may cause greater changes to the world than any other man-made factor in our environment. On the other hand, this may be a completely false alarm".

Twenty years ago, most members of the 'scientific establishment' thought it was a false alarm. It was generally believed that temperature fluctuations such as had occurred over the past few thousands of years would be much greater than any induced by man-made atmospheric changes.

Today the situation is quite different. There is general agreement that *some* effect will be manifested if greenhouse gases, particularly CO<sub>2</sub> and methane, continue to increase. And, as John Maddox (*Nature* 334, 9; 1988) concludes, governments must decide what to do when the effect is palpable. But we still do not know when this will happen, and what the magnitude of the effect will be.

I see in the press that there is now a

worry that the Thames barrier will be insufficiently high to cope. Some years ago, when this work was nearing completion, I visited the site with the Parliamentary and Scientific Committee. After our inspection, we had a question-and-answer session. I asked whether the possible effects of a rise of ocean levels due to greenhouse gases had been taken into consideration in planning the barrier. I was assured that it had. We now learn that it had not, which should be a lesson to our planners and civil engineers.

KENNETH MELLANBY

*38 Warkworth Street,  
Cambridge CB1 1EG, UK*

## Magic results

SIR—I was intrigued by your use of a professional magician in the inquiry into the conduct of experiments reported by Davenas *et al.* (*Nature* 333, 816–818; 1988), indicating biological activity of antibody solutions containing no antibody molecules. It is true that the results reported contradict accepted ideas, and scientists require much more persuading before believing them. But error and even fraud are known to exist in other much more conventional fields. This is especially true in medical research and biology, where there are many pressures. Even in physics, however, the desire to win the Nobel prize may lead to doubtful interpretations being suggested, as now appears to be the case for the claimed 'fifth force'. I suggest therefore that professional magicians be employed to verify all experimental results reported in *Nature*. This would not only lead to equal treatment of different scientists but would also give regular employment to magicians, at present dependent on the economics of show business.

MICHAEL FRIEDJUNG

*CNRS Institut d'Astrophysique,  
98 Bis, Boulevard Arago,  
75014 Paris,  
France*

## Men only?

SIR—The review by Stuart Sutherland of *A Passion for Science* (*Nature* 334, 112; 1988) emphasizes that scientists are tempted to steal colleagues' wives and pinch *au pairs'* bottoms. As well as musing over the changing colour of women's stockings (stockings?), he could perhaps have assessed whether a suitable subtitle might be *The Joy of Sexism*.

PAMELA V. TAYLOR

*Department of Obstetrics and  
Gynaecology,  
University of Leeds,  
'D' Floor, Clarendon Wing,  
Belmont Grove,  
Leeds LS2 9NS, UK*



# Reading tea-leaves in the sky

*Cosmologists spend much of their time poring over maps of the distribution of galaxies. A new way of searching for patterns may help them uncover the past, if not predict the future.*

THE ancient astronomers who dreamed up the constellations found deities, heroes and animals in the sky because those were the creatures they assumed to inhabit the heavens. Modern cosmologists have been known to do the same. Those whose cosmogonical theories imply that galaxies congregate in giant clusters, with huge empty regions in between, can find in the distribution of galaxies on the sky the patterns they seek; those with other theories argue that any apparent large-scale structures, either clusters or voids, are simply the statistical extremes of an underlying homogeneous population. Even worse, the human eye is notorious for finding orderly patterns in the most random of data, and when some cosmologists declare that galaxies are 'obviously' arranged in filaments or knots or loops, like the fabric of a coarse-meshed net flung on the sky, others have countered that the whole thing is a web of delusion invented by hopeful minds.

There is only one Universe — at least only one we can see — and cosmologists have never been able to do true comparison shopping among their models and theories. When observations can be made quantitative, predictions can be compared with measurements, but when, as in the case of large-structure, the reality of filaments and voids is a matter for subjective debate, models of the Universe can likewise be judged only by the same unsatisfactory standards.

Advances in observational technology have allowed astronomers to compile large catalogues of galaxies, listing their position, brightness and more recently their redshift. At the same time, supercomputers permit simulation of the motion and gravitational interaction of thousands of galaxies over billions of years. Cosmologists thus have at their disposal three-dimensional maps of the real Universe and computer simulations of comparable complexity. The weak spot has been in making objective comparisons between the two, and this has hindered attempts to winnow out good from bad in the ingredients that go into the models.

Until about the late 1960s, the statement that galaxies were smoothly distributed on sufficiently large scales was enough for most purposes. But as cosmologists began to grapple with the problems of how and when galaxies are made, more sophisticated use of the data was needed.

Statistical analysis of the galaxy distribution

was introduced by P.J.E. Peebles in the early 1970s. His most powerful device was the two-point correlation function, which is the distribution of the distance between every galaxy in a survey and its nearest neighbour. Peebles and his students were able to prove quantitatively that galaxies tend to cluster: nearest-neighbour distances are systematically smaller than would be expected for points scattered at random on the sky.

Galaxy formation theories stood or fell by their ability to reproduce the observed two-point correlation function, but because this statistical measure was simple and powerful there was a tendency to ignore those features which it could not accommodate, and to imagine that it contained essentially all there was to know about the distribution of galaxies.

But there was unease too. Pictures of galaxy distributions generated by computers from simple theories passed the statistical tests, but somehow did not look like the real sky. The models looked too uniform, and Soviet cosmologists in particular declared that real galaxies were assembled in filaments, sheets and knots, to which the correlation functions were insensitive because of the way they averaged over all galaxies and in all directions.

Efforts to quantify the large-scale structure, or to prove its absence, go back many years. Fractal geometries became popular for a while, but application of fractals to galaxies seemed to do little but express the correlation function in a different language. Percolation theory was used: circles of increasing size can be inscribed around every galaxy, with the idea that at some characteristic radius all the disconnected islands join into a single connected structure (the pattern is said to 'percolate'). This radius ought to be smaller if galaxies are concentrated in linear patterns rather than scattered at random, and this indeed seemed to be the case. But the percolation radius is not easily incorporated into theoretical analyses, and although such ideas began to convince people that filaments of galaxies are probably real, there was little progress in making useful comparisons of models with the data.

With the idea of filaments and voids more or less accepted but not adequately characterized, cosmology conferences were entertained by debates over whether the Universe is more like a Swiss cheese or a sponge (in a Swiss cheese, the holes are

isolated, but in a sponge, both material and empty space form connected, complementary structures: soak a sponge in concrete, dissolve away the foam, and another sponge is left).

In the current issue of *Astrophysical Journal* (331, L63; 1988), Suketu Bhavsar of the University of Kentucky and Nigel Ling of the University of Sussex say they have put an end to this debate with an algorithm which, applied to the three-dimensional galaxy survey from the Harvard-Smithsonian Center for Astrophysics, demonstrates that filaments really exist; that is, compared to the chance alignments that would be expected in a random distribution, there are more of them, and they are longer.

In the complete catalogue of galaxies, Bhavsar and Ling construct the 'minimal spanning tree', a unique network of lines which joins every galaxy in such a way that the total length of all the connecting lines is minimized. Two further operations are performed. The minimal spanning tree is 'pruned', by removing all branches which have fewer than some chosen number of galaxies on them, and the network is 'separated' into a small number of isolated networks by removing all links greater than some chosen length.

The purpose of these two procedures is to give prominence to filamentary structures, but it remains to be shown that the filaments thus revealed are 'real'. Bhavsar and Ling accomplish this demonstration by dividing the complete data set into smaller subunits, and randomly reassembling the subunits into a fake galaxy catalogue. If the filaments found in the original data were simply chance alignments, the randomization should remove them but would replace them in the fake data by a comparable selection of filaments. Instead, shuffling the subunits eliminates the filaments but does not produce new ones, implying that they were a special property of the real galaxy distribution, not an artefact of random data.

Variation of the lengths that define pruning and separation in the minimal spanning tree leads to quantitative measures of the distribution of filament lengths, which should allow statistical comparisons to be made between theory and reality. Then one can hope to decide whether filaments, if not galaxies, are randomly distributed, or whether yet deeper structures are waiting to be discerned.

David Lindley

## Oxide superconductors

## Electronic structure revealed

Ned Stoffel

To elucidate the mechanism responsible for superconductivity in the high-temperature oxide superconductors, the electronic structure of the materials must first be understood. Superconductivity is believed to involve the creation of pairs of charge carriers (Cooper pairs) through a subtle rearrangement of the conduction electrons near the Fermi energy  $E_F$  (the energy of the highest filled electronic states). The character of the Fermi-level states and the interaction which binds them into Cooper pairs are currently the subject of intense debate. On page 691 of this issue<sup>1</sup>, Takahashi *et al.* report a spectroscopic study of the 85-K superconductor  $\text{Bi}_2\text{CaSr}_2\text{Cu}_2\text{O}_8$  which reveals the nature of the electron states with unprecedented detail. By showing that the states involved in superconductivity display an energy resonance which is characteristic of oxygen atoms, they provide vital clues for resolving the mechanism.

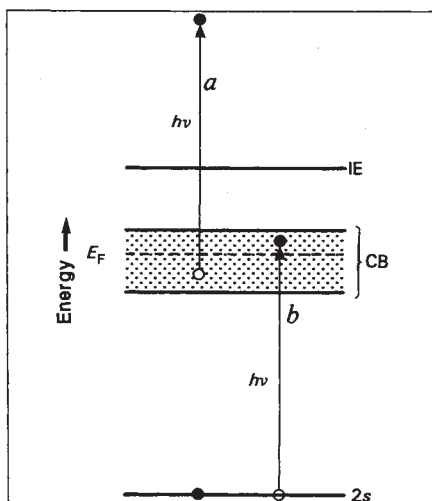
Angle-resolved photoemission, the technique Takahashi *et al.* employ, is used extensively to study the valence electronic structure of semiconductors and metals. A clean, atomically ordered sample surface, prepared in a vacuum, is illuminated with monochromatic photons. The emitted photoelectrons are collected by a movable detector which simultaneously determines their energy and momentum. From these quantities, the initial energy and 'crystal' momentum of the electron in the solid can be inferred assuming conservation of energy and momentum in the photoemission process.

In principle, the energy-momentum distribution (band structure) of the electrons in the new superconductors could be mapped in this way. In practice, however, this goal has been difficult to attain because of the low photoemission signal at the Fermi level  $E_F$  and because of difficulties associated with the preparation of an ordered, chemically stable surface. Takahashi *et al.* have succeeded not only in measuring several electron bands (the sets of allowed energies), one of which seems to cross the Fermi level; they can also demonstrate an enhancement in the photoemission probability for the band at  $E_F$  for photon energies near 18 eV.

This energy corresponds to the absorption edge of the first inner-shell excitation in oxygen (see figure). This effect, related to the autoionization resonance seen in atoms, suggests that the character of the states at the Fermi energy is determined principally by the oxygen atoms in the crystal. The result complements earlier experiments<sup>2</sup> which associated the

character of oxygen 2p-electron orbitals with the empty electronic levels immediately above  $E_F$ . The electronic states in the 3d-shell of copper atoms also span the Fermi level, but the Fermi level states do not show a resonance effect at the Cu-3p inner-shell absorption threshold.

The band dispersions (energy-momentum relations) detected by Takahashi *et al.* in  $\text{Bi}_2\text{CaSr}_2\text{Cu}_2\text{O}_8$  encompass only a fraction of the bandwidth predicted by



Takahashi *et al.*<sup>1</sup> use angle-resolved photoemission spectroscopy, in which the direction and energy are analysed of electrons ejected from the sample by incident light, to reveal the electronic structure of the superconductor  $\text{Bi}_2\text{CaSr}_2\text{Cu}_2\text{O}_8$ . The direct photoemission of electrons from near the Fermi level  $E_F$  (a) is resonantly enhanced when the incident photon energy  $h\nu$  coincides with the energy of an allowed 'inner-shell' transition (b). The enhancement can be used to determine the atomic origin of the direct photoelectron. Oxygen atomic level 2s; IE, ionization energy; CB, conduction band.

electron band theory. Our earlier attempt<sup>3</sup> to apply angle-resolved photoemission to  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  crystals did not reveal any clear indications of the dispersion. In both cases, the suppression of band dispersion could be attributed to the effects of strong electron-electron interactions in these materials. Abundant evidence of these effects within the copper valence electrons is already available from earlier spectroscopic experiments<sup>4</sup>.

Electron band theory is extremely successful in explaining the transport properties and optical response of metals and semiconductors. But it is most appropriate for materials, such as the alkali metals, in which the electrons can be treated as virtually independent particles moving in the time-average potential of the crystal lattice. In metals, the Coulomb

repulsion between the electrons is partially screened by the dielectric response, but the residual repulsive forces cause the electrons to move in correlated orbits to maximize their separation and to minimize their Coulomb-interaction energy.

The effects of correlation can be included in the calculations if they are a small perturbation of the crystal potential. But in ionic materials, including many metal oxides, the valence-electron orbits have small overlaps and are rather localized. The poorly screened Coulomb repulsion between two electrons on an atomic site can become comparable to the bandwidth, precluding an accurate description by conventional band theory.

The effects of large correlation energies on electronic structure can be calculated with reasonable accuracy for atoms and small clusters, but similar calculations in solids require significant approximations. The Mott-Hubbard theory, which treats the Coulomb repulsion as the dominant energy, is a useful approach for dealing with highly correlated systems and explains why some materials, including the undoped superconducting compounds, are insulators when simple band theory predicts that they should be metallic. Anderson's resonating-valence-bond (RVB) model<sup>5</sup> starts from the Mott-Hubbard description and has been used successfully to predict some of the unusual magnetic behaviour of the superconducting compounds. The model is not intended to be a complete description of the electronic structure, however, and some energy terms which could be significant are not included. In fact, in many models of this type, the basic energy parameters are unknown and it is doubtful that they can be calculated from first principles.

The presence of oxygen states near the Fermi energy adds significantly to the complexity of the Hubbard-type models. Additional parameters are needed to describe the effects of interatomic charge transfer, magnetic fluctuations, and inter-site Coulomb repulsion between copper and oxygen ions. The physical properties derived from the models depend on the relative sizes of these contributions<sup>6</sup>, and to some extent this has led to a remarkable variety of proposed explanations for high-temperature superconductivity. Clearly, the basic electronic structure needs to be established with more precision to clarify the theoretical situation, and this will require further experimental measurements on single crystals. □

1. Takahashi, T. *et al.* *Nature* **334**, 691-692 (1988).
2. Nücker, N., Fink, J., Fuggle, P.J. & Temmerman, W.M. *Phys. Rev. B* **37**, 5158-5163 (1988).
3. Stoffel, N.G. *et al.* *Phys. Rev. B* **37**, 7952-7955 (1988).
4. Fujimori, A., Takayama-Murochiri, E., Yuchida, Y. & Okai, B. *Phys. Rev. B* **35**, 8814-8817 (1987).
5. Anderson, P.W. *Science* **235**, 1196-1198 (1987).
6. Zaanen, J., Sawatzky, G.A. & Allen, J.W. *Phys. Rev. Lett.* **55**, 418-421 (1985).

Ned Stoffel is at Bellcore, 331 Newman Springs Road, Red Bank, New Jersey 07701, USA.



## Cell physiology

## A real receptor-operated calcium channel?

Timothy J. Rink

SOME of the mechanisms by which activator calcium enters stimulated cells are increasingly well understood as the structure and function of voltage-gated calcium channels are elucidated; others remain more elusive. The existence of receptor-operated calcium channels (ROCCs) was suggested on the basis of smooth muscle responses which seemed to depend on calcium entry, when membrane potential did not change, or when voltage-gated calcium channels were blocked by specific 'calcium antagonists' (ref. 1). ROCCs have also been postulated to explain calcium entry into cells that are not electrically excitable, as in thrombin-stimulated blood platelets. Now Zschauer and colleagues report on page 703 of this issue<sup>2</sup> the demonstration of single channels with the expected properties of ROCCs.

Although ROCCs are an attractive way to explain phenomena such as those mentioned above, it was not clear whether they could be explained by other means, for example, other transduction processes requiring external  $\text{Ca}^{2+}$  to generate some other intracellular message, or voltage-gated channels not affected by the calcium antagonists. More recently, it has been shown in cells that seem to lack voltage-gated calcium entry, including platelets, parotid-gland and endothelial cells<sup>3-5</sup>, that receptor-mediated elevation of cytosolic calcium is greater and more prolonged in the presence of external calcium than in calcium-free medium. This suggests, but does not prove<sup>6</sup>, that receptor occupation stimulates calcium entry through some form of ROCC.

Additional support comes from the demonstration<sup>7</sup> of stimulated uptake of radio-labelled  $^{45}\text{Ca}^{2+}$  and of stimulated entry<sup>8</sup> of another divalent cation,  $\text{Mn}^{2+}$ . There are many ways that receptor occupation could be coupled to calcium entry, including a direct effect on a subunit of the channel, analogous to the nicotinic acetylcholine receptor-channel complex; a G-protein; the action of  $\text{Ca}^{2+}$  itself; inositol phosphates; or the action of a protein kinase. Moreover, there are various possible mechanisms for the calcium transport, including a transmembrane channel, or pore, across the plasma membrane; an ionic exchanger; or a specialized pathway direct from the external medium to the interior of an internal organelle (see ref. 5). For many, the only convincing evidence of ROCCs would be the demonstration of calcium currents through single

channels that show no voltage dependence in opening and closing.

This type of evidence has been difficult to obtain. One of the first convincing examples was the ATP channel examined by patch-clamp in isolated smooth muscle cells, recently reported by Benham and Tsien<sup>8</sup>, which is fairly selective for calcium over sodium. In T lymphocytes<sup>9</sup>, and now in mast cells<sup>10</sup>, evidence has been presented that intracellular inositol trisphosphate ( $\text{InsP}_3$ ) can promote calcium-channel activity. A role for inositol tetrakisphosphate ( $\text{InsP}_4$ ) in calcium entry remains controversial<sup>10</sup>.

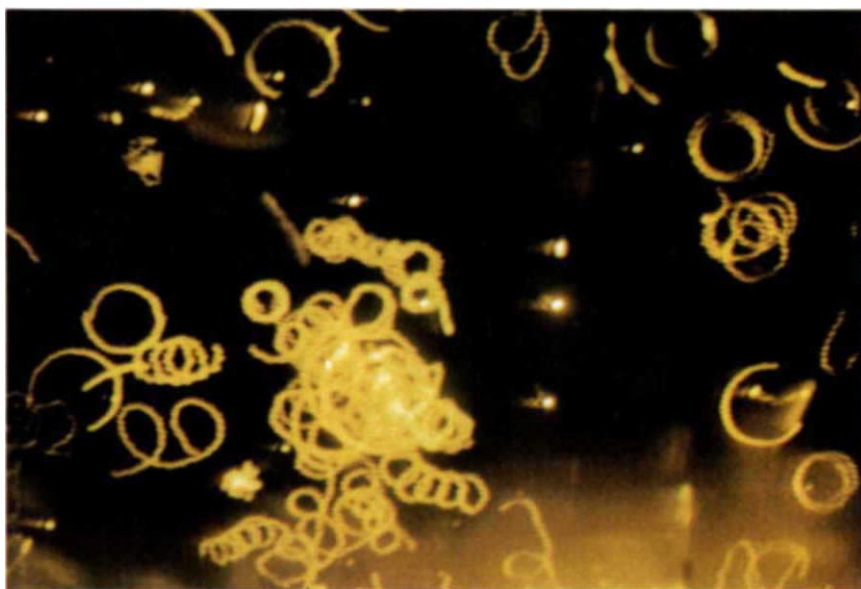
Perhaps the most striking demonstration of a ROCC is that reported by Zschauer *et al.* in this issue<sup>1</sup>. The authors have incorporated membrane vesicles from resting and thrombin-stimulated platelets into planar bilayers, and observe single channels 30-fold selective for the divalent cation  $\text{Ba}^{2+}$  over  $\text{Na}^+$  in vesicle-stimulated cells. The conductance, with

symmetrical solution containing 150 mM  $\text{Ba}^{2+}$  (a commonly used surrogate for  $\text{Ca}^{2+}$  in such experiments) is about 10 pS in the concentration range reported for voltage-gated calcium channels in similar conditions. The membrane fraction clearly contains plasma membrane, but it cannot be excluded that the channel activity comes from contaminating intracellular organelles.

The opening and closing kinetics of the channels derived from platelet vesicles are not influenced by membrane potential or by 10  $\mu\text{M}$  nisoldipine, a potent dihydropyridine antagonist of 'L-type' voltage-gated channels. But these channels are blocked by  $\text{Ni}^{2+}$ , which is known to block thrombin-stimulated calcium entry into platelets<sup>5</sup>. Thus, the results of Zschauer and co-workers do point to real ROCCs in thrombin-stimulated platelets, as we had earlier proposed from more indirect evidence<sup>3,6</sup>.

It is to be hoped that Zschauer and colleagues go on to show that  $\text{Ca}^{2+}$  itself is permeable through these channels, and to examine the effects of a different class of calcium antagonists. An intriguing and surprising point emerging from their data is that the thrombin-induced channels survive the cell-fractionation procedure. This result suggests that either some covalent modification, perhaps phos-

## Chemoattractant for sea urchin sperm



THE egg peptide resact ( $\text{Cys-Val-Thr-Gly-Ala-Pro-Gly-Cys-Val-Gly-Gly-Gly-Arg-LeuNH}_2$ ) represents a species-specific chemoattractant for sea urchin spermatozoa and is the first natural animal sperm chemoattractant identified. Here, approximately 1 picolitre of a 100-nM resact solution was added to a drop of spermatozoa; within seconds the cells congregate at the point of injection. The tracks represent the movement of the sperm cell over a 2-second exposure. Resact binds to a spermatozoan plasma membrane protein identified as the enzyme guanylate cyclase, whose primary structure is presented by D. V. Goeddel and collaborators on page 708 of this issue. The spiral pattern in the centre of the picture describes the path the sperm cell takes when moving in the gradient. Before it detects the gradient, the sperm cell swims in a circle, seen here on the outside of the picture. (Photo, L.J. Dangott.) □

phorylation, of a pre-existing membrane component occurs, or that additional membrane containing these channels is inserted into the plasma membrane. The minimum latency of 200 milliseconds, known from thrombin activation of calcium entry<sup>11</sup>, allows sufficient time for either. It is not clear if Zschauer and co-workers were lucky or judicious in their choice of thrombin as the stimulus because, unlike all other agonists we have looked at, thrombin promotes an influx of  $\text{Ca}^{2+}$  that lasts many minutes. It will be interesting to see if a rapidly desensitizing ligand, for instance ADP or platelet-activating factor, also induces these channels.

Let us hope that this demonstration of a 'real' ROCC will help to define the function and structural basis of this somewhat elusive phenomenon, although our

efforts are likely to be seriously retarded until high-affinity, specific ligands are discovered. □

1. Bolton, T.B. *Physiol. Rev.* **59**, 607–718 (1979).
2. Zschauer, A., van Breemen, C., Buhler, F.R. & Nelson, M.T. *Nature* **334**, 703–705 (1988).
3. Rink, T.J. & Hallam, T.J. *Trends biochem. Sci.* **9**, 215–219 (1984).
4. Merritt, J.E. & Rink, T.J. *J. biol. Chem.* **262**, 17362–17369 (1987).
5. Jacob, R., Hallam, T.J., Merritt, J.E. & Rink, T.J. *Nature* (in the press).
6. Hallam, T.J. & Rink, T.J. *FEBS Lett.* **186**, 175–179 (1985).
7. Massini, P. & Luscher, E.F. *Biochim. biophys. Acta* **436**, 652–663 (1976).
8. Benham, C.D. & Tsien, R.W. *Nature* **328**, 275–278 (1988).
9. Kuno, M. & Gardner, P. *Nature* **326**, 301–304 (1987).
10. Penner, R., Matthews, G. & Neher, E. *Nature* **334**, 499–504 (1988).
11. Sage, S.O. & Rink, T.J. *J. biol. Chem.* **262**, 16364–16369 (1987).

Timothy J. Rink is at Smith, Kline & French Research Ltd, The Frythe, Welwyn, Hertfordshire AL6 9AR, UK.

## Cosmology

# Submillimetre excess

Bernard Carr

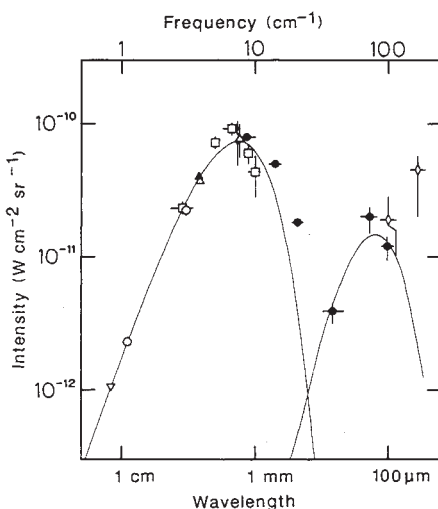
INCREASINGLY accurate measurements over a wide range of wavelengths have confirmed that the cosmic microwave background radiation has a perfect black-body spectrum, as required in the simplest Big Bang picture of the Universe, with a temperature of 2.7 K. But Matsumoto *et al.* now report<sup>1</sup> a significant distortion in the submillimetre waveband 400–700  $\mu\text{m}$  (wavelengths just short of the continuum maximum at 1,600  $\mu\text{m}$ ). This follows a rocket experiment by a team from Nagoya and Berkeley universities in February 1987. If real, this result is of crucial cosmological significance. Indeed, it has already prompted a flurry of theoretical papers, including one by the experimenters themselves<sup>2</sup>.

The microwave background, first discovered 23 years ago, shows that the Universe is gently glowing with radiation that is a relic from the Big Bang and has been cooled to 2.7 K by the expansion of the Universe. There have been earlier reports of distortions in the submillimetre region but these distortions have always been suspect, because of the uncertain error bars<sup>3</sup> or possible contamination from rocket exhaust<sup>4</sup>. The striking point about the Nagoya–Berkeley result, as shown in the figure, is that the error bars are so small.

The brightness is measured in six passbands by a liquid-helium-cooled radiometer calibrated intermittently throughout the flight using a black-body source. The measurement at the largest wavelength (1,160  $\mu\text{m}$ ) confirms the usual microwave temperature, and measurements at the three shortest wavelengths (262, 137 and 102  $\mu\text{m}$ ) are well explained

by interstellar dust emission. The surprising results are the intensities at the two intermediate wavelengths (709 and 481  $\mu\text{m}$ ) which are significantly in excess of the expected 2.7-K black-body flux, even though there is no obvious local source of radiation at these wavelengths.

Theorists envisage two types of explanation for the distortion. First, the microwave photons could have interacted with hot electrons while passing through ionized intergalactic gas. The thermal



Observed intensities of the cosmic microwave background including the new data of Matsumoto *et al.*<sup>1</sup> (solid circles). The solid lines represent the black-body component as calculated for a temperature of 2.74 K (left) and the emission spectrum of interstellar dust at 20 K (right). The measurements at 709 and 481  $\mu\text{m}$  clearly exceed the predicted background. (From ref. 1.)

*bremssstrahlung* (scattering) radiation from such gas has already been invoked to explain the background of hard X-rays. In this case, the microwave photons would gain energy from the electrons, with energy being depleted at wavelengths longer than the peak (the Rayleigh–Jeans region) and redeposited at wavelengths shorter than the peak (the Wein region). This ‘Compton distortion’ model only gives a good fit to the data if the temperature inferred from the Rayleigh–Jeans region is much lower than conventionally assumed<sup>2</sup>. There are also problems in finding a mechanism to heat the intergalactic gas to the high temperatures required. Stellar explosions could provide the necessary heat input<sup>5</sup> but only if one invokes a non-standard stellar mass function<sup>6</sup>. Otherwise the model needs exotic mechanisms such as superconducting cosmic strings (see the News and Views article by C. Hogan<sup>7</sup>).

The second possibility is that the submillimetre excess represents emission from cosmological dust, the dust itself having been heated by some radiation source<sup>8–10</sup>. The dust required either could be confined to galaxies (if galaxies form early enough to cover the sky, which is uncertain) or it could be a relatively smooth intergalactic component. Providing it is produced sufficiently early — the minimum redshift depends upon the type of dust and its abundance but is at least 10 — one would expect any pre-existing optical and ultraviolet radiation to be reprocessed into the submillimetre band. Indeed, the existence of a 400–700  $\mu\text{m}$  excess was predicted on this basis even before the Nagoya–Berkeley experiment<sup>8</sup>.

Although the dust model can fit the submillimetre spectrum rather well, it too faces severe energetic problems. Instead of heating up the intergalactic gas, as in the Compton model, one must heat up the dust, but the energy requirements are similar. This is reflected in the enormous energy density contained in the reported excess, which is equal to 10–20 per cent of the energy in the microwave background itself.

There are ways to provide the radiation energy required in the dust picture but they all tend to be fairly exotic<sup>8</sup>. If one invokes the radiation from stars, for example, they cannot have the same mass distribution as the stars in the solar neighbourhood because they would generate too many metals<sup>6</sup>. One would have to invoke ‘very massive objects’ (stars larger than 200 solar masses) because these collapse directly to black holes without any metal ejection. Such a hypothesis is not inconceivable — indeed the black hole remnants of very massive objects have already been postulated as candidates for the dark matter in galactic halos — but no such stars seem to be form-



ing at the present epoch.

Another possibility is that black-hole accretion could generate the required radiation. For example, there is a general consensus that supermassive black holes may reside in some galactic nuclei and, if they all went through a period of maximal accretion (so called Eddington-limit accretion) they could just about generate the required energy density. Alternatively, the radiation could be produced by decaying elementary particle relics of the Big Bang<sup>11</sup>. Such particles (or their decay products) are often invoked as candidates for dark matter and, with sufficient tuning of their mass and decay time, one could certainly explain the observed distortion. Indeed the possibility of such a distortion resulting from particle decays was predicted before the Nagoya-Berkeley results<sup>12</sup>.

Perhaps the most conservative explanation for the submillimetre excess would be to invoke the most prominent sources of far-infrared radiation at the present epoch, the starburst galaxies discovered by IRAS (the Infrared Astronomical Satellite). If one extrapolates the luminosity evolution associated with the IRAS galaxies to sufficiently high redshifts (itself rather speculative), one could in principle produce a background with the wavelength and energy density required. In this case, the submillimetre excess is coming from dust (as before) but from dust in the same galaxies where the radiation is produced. But a first estimate suggests that this model does not match the Nagoya-Berkeley data very well<sup>13</sup>.

Of course, the question remains of whether the excess is real. The detection of a cosmological background in any waveband is notoriously difficult and backgrounds in the infrared-millimetre range have appeared and then disappeared before. Fortunately, the matter is unlikely to be resolved until the launch of the Cosmic Background Explorer (COBE) satellite in 1989, leaving theorists with plenty of time in which to stretch their imaginations. □

1. Matsumoto, T. *et al. Astrophys. J.* **329**, 567-571 (1988).
2. Hayakawa, S. *et al. Pub. astr. Soc. Jap.* **39**, 941-948 (1987).
3. Woody, D.P. & Richards, P.L. *Astrophys. J.* **248**, 18-37 (1981).
4. Gush, H.P. *Phys. Rev. Lett.* **47**, 745-748 (1981).
5. Yoshioka, S. & Ikeuchi, S. *Astrophys. J.* **323**, L7-L11 (1987).
6. Lacey, C.G. & Field, G.B. *Astrophys. J.* **330**, L1-L4 (1988).
7. Hogan, C. *Nature* **326**, 742-743 (1987).
8. Bond, J.R., Carr, B.J. & Hogan, C.J. *Astrophys. J.* **306**, 428-450 (1986).
9. McDowell, J.C. *Mon. Not. R. astr. Soc.* **223**, 763-786 (1986).
10. Negroponte, J. *Mon. Not. R. astr. Soc.* **222**, 19-36 (1986).
11. Stebbins, A. & Silk, J. *Astrophys. J.* **300**, 1-19 (1986).
12. Kawasaki, M. & Sato, K. *Phys. Lett.* **B169**, 280-284 (1986).
13. Rowan-Robinson, M. & Carr, B.J. in *The Post-Recombination Universe* (eds Lasenby, A. & Kaiser, N.) (Reidel, Dordrecht, 1988).

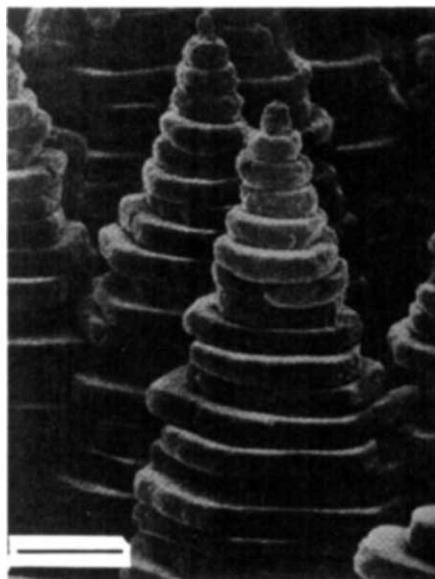
Bernard Carr is in the Astronomy Unit, Queen Mary College, London E1 4NS, UK.

## Biom mineralization

# Advances in mussel building

Paul Calvert

MINERALIZED biological tissues are very impressive composite materials. But because of the complexities of calcium and phosphate metabolism, there is no coherent picture of how mammals control mineralization to lay down the inorganic structures of bones and teeth. Calcium carbonate shells provide many parallels with bone formation and are more accessible to study. On page 692 of this issue<sup>1</sup>, S. Mann *et al.* report on the controlled crystallization of calcium carbonate on a



Growth surface of a nacreous layer of the mollusc *Monodonta labio*, showing aragonite tablets piled into conical stacks. Scale bar, 2.5  $\mu\text{m}$ . (From ref. 2.)

monolayer of stearic acid spread on the surface of a calcium bicarbonate solution. Stearic acid is a long-chain fatty acid and therefore the monolayer resembles the surface of a cell membrane. Mann *et al.* observe that vaterite forms rather than the normal stable polymorph, calcite.

Mollusc shells display many different arrangements of calcium carbonate, reflecting the mechanical constraints imposed by the life style of the species. One such arrangement is nacre. Nacreous shells provide a most attractive system for study, with plate-like crystals of aragonite about 0.5  $\mu\text{m}$  thick arranged in layers (see figure). These can be either staggered, with edges in one layer under centres in the next like bricks in a wall, or stacked. The crystals grow within pre-existing organic compartments which define their nucleation site, orientation and final size<sup>2</sup>. But why does aragonite form rather than the more stable calcite? What causes the crystals to nucleate at defined sites? And what causes them

to orient with their flat face down and, often, with their in-plane axes parallel?

Studies of extracts of mineralizing shells suggest that soluble, acidic, sulphated glycoproteins are responsible for the nucleation of calcium carbonate, but no one has yet managed to use them to form aragonite<sup>3,4</sup>. Weiner *et al.*<sup>5</sup>, following electron-diffraction studies of the *Nautilus* shell, argue that there is a relationship between the crystal axes within the planes and the underlying structure of chitin, a structural polysaccharide similar to cellulose found in crustacean shells, which somehow transmits its influence through the intervening layer of acidic protein. These studies are unsatisfactory because we do not really know the composition of the protein extracts and need some wholly synthetic, totally controlled systems for comparison.

There have been studies of crystallization on various surfaces such as glass, polystyrene and sulphonated polystyrene<sup>4,6</sup>. The surfaces do give rise to different forms and morphologies of calcium carbonate, but the surfaces of solids in liquids are usually poorly characterized and one can only speculate about the mechanism.

In their paper in this issue<sup>1</sup>, Mann *et al.* describe the nucleation of  $\text{CaCO}_3$  on monolayers of stearic acid spread on the surface of a solution of calcium bicarbonate which slowly loses  $\text{CO}_2$  and forms carbonate. They find that vaterite, a metastable polymorph of calcium carbonate, is nucleated with its (0001) face parallel to the surface. As with all 'Langmuir' films, the monolayer of stearic acid forms with the carboxylate ( $-\text{CO}_2^-$ ) ends in solution and the fatty chains in air. Presumably, the calcium ions in the solution attach themselves to the carboxylate as the first stage of crystallization. Mann *et al.* note that if the subsequent layer of carbonate is orientated in the same way as the carboxylate group, vaterite should be formed. The (0001) face of calcite has the carbonate ions lying parallel to the plane of the water surface (see their Fig. 3 on page 694). This evidently suggests a stereochemical mechanism for generating vaterite in preference to calcite. In the absence of the surface film, calcite forms on the walls and surface of the tank. As Mann *et al.* increase the surface pressure and impose more order on the films, nucleation becomes more rapid. Intermediate pressures give a few localized nucleation sites. Thus there are interesting parallels with nacreous shell growth.

Although Mann *et al.* can control the

structure and properties of the surface, they do not have complete control of the solution. It is surprisingly difficult to make crystallizing solutions of calcium carbonate with controlled supersaturation and pH. Mann *et al.* rely on the loss of carbon dioxide from a bicarbonate solution to increase slowly the supersaturation. This leaves questions about whether the surface concentration differs from the bulk and the actual supersaturation.

Mann *et al.* seem to suggest that crystal nucleation occurs heterogeneously on the surface layer by the build-up of ions on the calcium stearate surface. But Rieke, who has observed the growth of  $\text{CaCO}_3$  crystals on glass and polystyrene surfaces<sup>7</sup>, disagrees. The former generate calcite and the latter vaterite. Vertical surfaces grow fewer crystals than horizontal ones. Reike suggests that nucleation occurs homogeneously in solution, possibly to form an amorphous calcium carbonate. These

nuclei attach to the surfaces, which cause them to convert to one or the other form of the crystal that subsequently grows. This is intriguing, because there is evidence of amorphous precursors in bone growth and amorphous calcium carbonates have been postulated in some shell mineralization. □

1. Mann, S., Heywood, B.R., Rajam, S. & Birchall, J.D. *Nature* **334**, 692–695 (1988).
2. Nakahara, H. in *Biomaterialization and Biological Metal Accumulation* (eds Westbroek, P. & De Jong, E.W.) 225–230 (Reidel, Dordrecht, 1983).
3. Greenfield, E.M., Wilson, D.C. & Crenshaw, M.A. *Am. Zoologist* **24**, 925–932 (1984).
4. Addadi, L., Moradian, J., Shay, E., Maroudas, N.G. & Weiner, S. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2732 (1987).
5. Weiner, S., Talmon, Y. & Traub, W. *Int. J. Biol. Macromol.* **5**, 325–328 (1983).
6. Rieke, P.C. in *Atomic and Molecular Processing of Electronic and Ceramic Materials* (eds Aksay, I.A., McVay, G.L., Stoebe, G.L. & Wager, J.F.) 109–114 (Materials Research Society, Pittsburgh, 1988).
7. Rieke, P.C. *Langmuir* (in the press).

Paul Calvert is in the School of Chemistry and Molecular Sciences, University of Sussex, Brighton BN1 9QJ, UK.

## Immunology

# Fate of self-reactive B cells

Anthony L. DeFranco

A CENTRAL problem for the immune system is how to distinguish components of invading microorganisms from components of the body. In autoimmune disease, this discrimination breaks down, with severe pathological consequences. The paper by Goodnow *et al.* on page 676 of this issue<sup>1</sup> provides convincing evidence for the notion that the mechanisms for this discrimination are different in the T- and B-lymphocytes of the immune system.

Lymphocytes have receptors for antigens; each cell creates its own receptor by joining multiple, discrete segments of the genes encoding these receptors. In the process, many different receptors are generated, allowing the immune system to recognize an almost limitless number of distinct molecular structures. The antigen receptors on some lymphocytes recognize foreign antigens, whereas those on others recognize self components. Thus, the ability to discriminate self from non-self must be made by lymphocytes after the acquisition of antigen receptors. An important strategy that the immune system uses is to inactivate lymphocytes reactive to self antigens. But, as the receptors for any given antigen are typically present on a small fraction of the cells, it is difficult to observe directly what happens to lymphocytes that recognize self antigens, particularly with B lymphocytes, the cells that make antibodies.

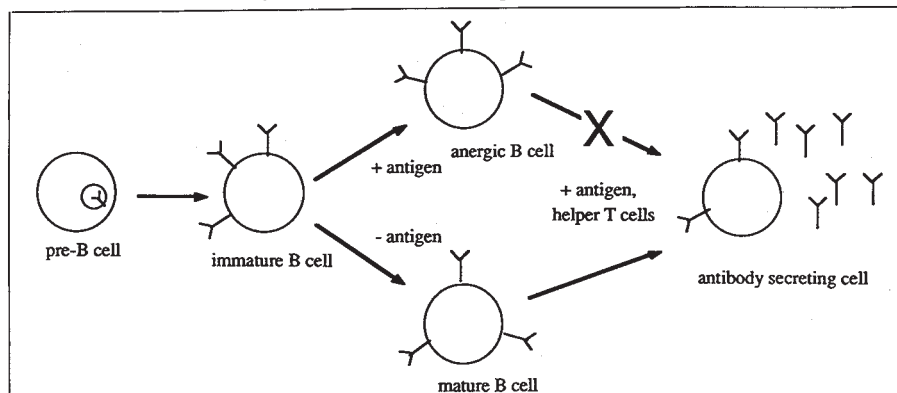
Goodnow and colleagues<sup>1</sup> have created transgenic mice that express a high-affinity antibody specific for hen-egg lysozyme as the antigen receptor on most

of their B lymphocytes. When lysozyme is made a self-antigen for these animals, by introduction of a lysozyme transgene, lysozyme-specific B cells are found to be present in high numbers, but are functionally inactive. This is in contrast to self-reactive T lymphocytes, which have recently been observed to die rapidly before exit from the thymus<sup>2–4</sup>. Thus, the mechanisms of self–non-self discrimination appear to differ somewhat between B lymphocytes and T lymphocytes.

Immunological tolerance of self components is a complicated phenomenon, likely to be mediated by several distinct

mechanisms. As stated above, an important mechanism for establishing tolerance is the purging of autoreactive lymphocytes following their interaction with antigen at an immature stage of their development<sup>2–6</sup>. Lymphocytes develop from haematopoietic stem cells late in the development of the embryo and on a continuing basis after birth. Once the developing lymphocyte successfully rearranges the genes it needs to make a functional antigen receptor, the cell arrives at an immature stage. At this point, if the cell contacts an antigen, it becomes inactivated. On the other hand, if the immature lymphocyte does not contact an antigen, it progresses to the fully functional mature stage that can make an immune response on subsequent challenge by antigen (see figure). The immune system takes advantage of the fact that most self-components are present continuously. It sees infectious agents that appear some time after birth as foreign, because lymphocytes specific for these agents were able to mature before infection.

Functional inactivation of immature B cells by antigen was demonstrated and characterized in considerable detail by Nossal, Klinman and collaborators<sup>5</sup>, as well as by others, in the 1970s. In general, these studies followed the fate of B cells by using the functional ability of B cells to make antibodies. Contact of the developing B cell with antigen at a critical period prevents it from responding later, but the actual fate of the tolerized B cell was unknown. This phenomenon was called<sup>6</sup> 'clonal deletion', implying that the inactivated B cells die. Later, functionally inactivated B cells were found to be viable and capable of binding antigen, suggesting that 'clonal anergy' was a better description of the fate of self-reactive B cells. Although elegant and informative, these experiments had the limitation that it was



Development of B lymphocytes. The precursor of the B cell, the pre-B cell, expresses the heavy chain of immunoglobulins but neither of the two conventional light-chain types. The heavy chain is expressed intracellularly, in combination with an alternative form of light chain ( $\omega$ )<sup>13</sup>. Once a light-chain gene is successfully rearranged to give rise to an encoded light chain, the cell expresses IgM on its surface and is an immature B cell. If this cell contacts antigen (top), it is functionally inactivated (anergic B cell), whereas if it does not contact antigen in the brief period following membrane IgM expression, then it can progress to the mature B cell stage (bottom). Later the antigen, in conjunction with helper T cells, may appear and trigger an antibody response of the mature B cell, but not of the anergic B cell. Y, Immunoglobulin. □



difficult to observe what happened to the self-reactive B cells, especially *in vivo*.

This problem is neatly solved by Goodnow *et al.*<sup>1</sup>, who describe the introduction of functionally rearranged immunoglobulin genes into the germ line of mice, thereby creating mice in which almost all B cells bind with high affinity to one antigen, hen-egg lysozyme. As expected, these mice have large antibody responses to lysozyme. Because the lysozyme-specific B cells are so abundant, it is fairly easy to see what happens to them under various circumstances. Next, Goodnow *et al.* introduced the gene encoding lysozyme into the germ line of other mice, yielding mice for which lysozyme is a tolerated self-component. The lack of reactivity to lysozyme is manifest in both the B- and the T-cell populations. Crossing these transgenic mice yields animals expressing both lysozyme and anti-lysozyme transgenes. Many lysozyme-specific B cells are found in the spleens and lymph nodes of these mice, although their number seems to decline with age.

Despite the many antigen-specific B cells, the animals fail to make antibody responses to lysozyme. To test the functional capability of the B cells from the doubly transgenic mice, Goodnow *et al.* transferred small numbers of spleen cells from these mice to a non-transgenic host mouse containing lysozyme-specific helper T cells. In this environment, the lysozyme-specific B cells still fail to make an antibody response. The authors argue that the inability to respond is a property of the B cells and not of another transferred cell, such as a suppressor T cell, because only  $10^5$  cells were transferred. This is a reasonable, but not airtight, argument. In any case, the B cells were not clonally deleted, as occurs when immature T cells contact antigen<sup>2,4</sup>, but rather were rendered anergic. This supports earlier results from B lymphocytes *in vitro*<sup>5</sup>, and indicates that functional inactivation of immature lymphocytes by antigen is not identical for B cells and for T cells.

One aspect of these studies that needs to be clarified is the form of the antigen that induces B-cell tolerance. Studies of signal transduction by membrane immunoglobulins IgM and IgD demonstrate that crosslinking of B-cell antigen receptors is necessary to trigger the phosphoinositide signalling pathway<sup>7,8</sup>. Lysozyme, a monomeric protein in solution, should not be able to crosslink membrane immunoglobulin molecules on the surface of B cells and hence should not be able to trigger receptor signalling. But in addition to being present in serum, lysozyme is also bound to extracellular matrix components (C. Goodnow, personal communication), which could provide the crosslinking form of antigen that should be necessary to induce B-cell tolerance. (Do truly mono-

meric self proteins induce anergy of B cells that recognize them? Quite possibly not. In these cases, the B cells probably fail to make antibody responses because they lack antigen-specific helper T cells.)

The experiments of Goodnow *et al.*<sup>1</sup> demonstrate that a self antigen can induce anergy of antigen-specific B cells. Similarly, the recent work using T cells<sup>2,4</sup> demonstrates that antigen contact by immature T cells results in rapid clonal deletion. Thus, we now have a much clearer picture of the mechanisms of antigen-induced inactivation of immature lymphocytes. Nevertheless, we know much too little about the rules under which these mechanisms operate. Self-reactive B lymphocytes in normal individuals without evidence of autoimmunity, for example, can be induced to produce antibody<sup>9-11</sup>. The responsiveness of these self-reactive B cells has led some to question whether functional inactivation of immature B cells by antigen is a sufficiently efficient phenomenon to contribute to tolerance. The work of Goodnow *et al.*<sup>1</sup> argues strongly that it is, at least under some circumstances.

So how do the observed self-reactive B lymphocytes avoid or exit from clonal anergy? Are they cells that bind the self component with too low an affinity to

trigger clonal anergy<sup>12</sup>? Could they be the result of somatic mutation of their antigen-receptor genes? Another problem involves the accessibility of self antigens to the milieu of developing lymphocytes. Many self antigens may be sequestered inside cells or in particular organs, away from the site of B-cell development in the bone marrow and from T-cell development in the thymus. Do these antigens fail to induce inactivation of immature lymphocytes and, if so, how does the immune system learn that they are self? The transgenic mouse system seems to be ideally suited to answer these questions. □

1. Goodnow, C.C. *et al.* *Nature* **334**, 676-682 (1988).
2. Kappler, J.W., Roehm, N. & Marrack, P. *Cell* **49**, 273-280 (1987).
3. Kappler, J.W. *et al.* *Nature* **332**, 35-40 (1988).
4. Macdonald, H.R. *et al.* *Nature* **332**, 40-45 (1988).
5. Nossal, G.J.V. *A. Rev. Immun.* **1**, 33-62 (1983).
6. Burnet, F.M. *The Clonal Selection Theory of Acquired Immunity* (Cambridge University Press, London, 1959).
7. Cambier, J.C. & Ransom, J.T. *A. Rev. Immun.* **5**, 175-200 (1987).
8. DeFranco, A.L. *A. Rev. Cell Biol.* **3**, 143-178 (1987).
9. Weigle, W.O. *J. exp. Med.* **121**, 289-307 (1965).
10. Primi, D. *et al.* *J. exp. Med.* **145**, 21-30 (1977).
11. Harris, D.E., Cairns, L., Rosen, F.S. & Borel, Y. *J. exp. Med.* **156**, 567-584 (1982).
12. Riley, R.L. & Klinman, N.R. *J. Immun.* **136**, 3147 (1986).
13. Pillai, S. & Baltimore, D. *Nature* **329**, 172-174 (1987).

Anthony L. DeFranco is in the Department of Microbiology and Immunology, University of California, San Francisco, California 94143-0552, USA.

## Protein cytoskeleton

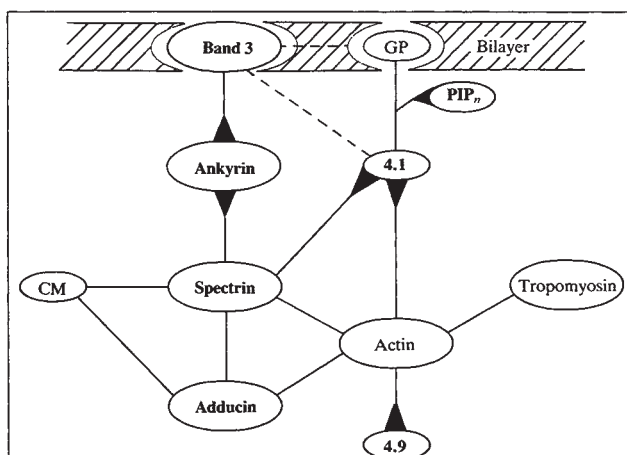
# Functional or futile phosphorus?

Lars Backman

THE human red-cell membrane is a complex piece of machinery that burns a great deal of metabolic energy to no discernible purpose. The cytoplasmic ATP pool is of course needed to drive the pumps that keep out unwanted ions, but much of the energy goes into an unceasing turnover of phosphoryl groups on the proteins of the membrane and its protein skeleton, and also on the inositol lipids. These processes are catalysed by a comprehensive range of kinases and phosphatases, both cytoplasmic and membrane-bound. There is no dearth of theories but no conclusive evidence that all this activity accomplishes anything at all for the cell, like preserving its discoid shape (which is certainly lost if the ATP is spent). On page 718 of this issue<sup>1</sup>, Branton and colleagues describe a mystifying new phenomenon of phosphorylation to leave students of the red cell and of actin-binding proteins better informed if no wiser.

Undoubtedly, the various phosphorylations in the red-cell membrane do affect the interactions between the different proteins *in vitro*. The interactions that have been recognized, and can be taken to prevail in the cell, are summarized in the figure, which also indicates points at

which phosphorylation or dephosphorylation could perturb the system. Most recently, for example, Morrow and colleagues<sup>2</sup> have shown that the phosphorylation of ankyrin, the protein that binds the membrane to the membrane skeleton by attaching at one end to the main transmembrane protein, band 3, and at the other to spectrin, alters the relative strength of its interactions. Ankyrin, according to Cianci *et al.*<sup>2</sup>, binds less strongly to the basic structural unit of spectrin, the  $\alpha\beta$ -dimer, than to the tetramers (and higher forms) that predominate in the cell. This preference for the associated state is enhanced by the attachment of band 3 (or at least the appropriate proteolytic fragment of this molecule) to the ankyrin. The magnitude of the enhancement is reduced when the ankyrin is phosphorylated. Earlier, the work of Tao's laboratory<sup>3</sup> had shown that phosphorylation of ankyrin also weakens its interaction with band 3 (whereas phosphorylation of the band 3 has no effect). If these phenomena operate *in situ* — and it has to be said that the effect described by Cianci *et al.*<sup>2</sup> is by no means large *in vitro* — the fine tuning of this system of interactions must be discouragingly complex. ▶



Schematic illustration of reported protein interactions in the red-cell membrane and in the skeleton. Proteins known to be phosphorylated are in bold text; arrowheads indicate interactions affected by phosphorylation. PIP<sub>n</sub>, phosphatidylinositol phosphates; CM, calmodulin; GP, glycophorin.

Now, Branton and co-workers<sup>1</sup> have hit on something much more like an all-or-none effect of a phosphorylation. But this result too has elusive features. Branton and collaborators study protein 4.9, an actin-binding and bundling protein of inscrutable, even paradoxical function. In solution it exists as a trimer, containing two structurally similar subunits of relative molecular mass 48,000 (48K) and 52K. As the subunits are present in a ratio of 3:1, some of the trimers must contain only the 48K subunit. Both subunits are substrates for cyclic AMP-dependent and -independent kinases. The phosphorylation, it transpires, inhibits actin bundling and at a concentration of two phosphoryl groups per 48K subunit, annihilates it. The effect is reversible, for on dephosphorylation with alkaline phosphatase the activity returns. (Phosphorylation by protein kinase C does not affect bundling, but as only one mole of phosphorus is incorporated it is unclear whether this is because the kinase acts at different sites or causes too low a degree of phosphorylation.)

When bundling is maximal (dephosphorylated protein) a 4.9 trimer binds on average to 11–13 actin subunits (about one turn of the actin helix). The inactive, phosphorylated form also binds to actin filaments but at a concentration of about one trimer per 23 subunits (a strange number), and without effecting any bundling. Herein lies a mystery: here we have a protein with three effectively (or in some trimers, actually) identical actin-binding subunits, yet once it is phosphorylated, it ceases to crosslink actin but still binds. The multivalency that is implied by the bundling activity is not then after all a function of the subunit structure. Moreover, Branton and colleagues find that dephosphorylated monomeric protein 4.9, generated by addition of glycerol, retains its bundling activity in contrast to the phosphorylated subunit. Either, then, phosphorylation knocks out a binding site,

or it distorts the geometry such that the sites cannot bind two filaments. But why in this case can a symmetrical trimer not make cross-links when there is still one working site per subunit?

And what is the protein doing there anyway? By radioimmunoassay, Branton and colleagues<sup>1</sup> find 40,000 copies of 4.9 trimer per cell, which corresponds closely enough to one for each of the actin protofilaments that make up the network junctions. The average separation of these filaments in a regu-

lar lattice, such as the membrane skeleton appears to be<sup>2</sup>, is far too large to make bundling a possibility. One might with an effort envisage a local, perhaps transient association of protofilaments if the lattice junctions were free to float in the membrane, but there is no evidence for such freedom. The complexity of the system is increased even further when adducin is considered. This calmodulin-binding protein, which, like protein 4.9, is phosphorylated by protein kinase C, also has actin-bundling activity<sup>3</sup>.

Of the other phosphorylations shown in the figure, that of 4.1 may be important, for it affects its interaction with spectrin and actin to form a prototypic skeleton complex *in vitro*<sup>4</sup>. But a causal link between phosphorylation and cell functions (in particular shape control) has been established only for the phosphoinositides<sup>5</sup>. One possibility is that phosphorylation-regulated interactions are important at the time of synthesis and assembly of the membrane skeleton and/or cytoplasmic cytoskeleton in the erythroid precursors, and that some of the components simply remain in the mature cells as now useless relics of their early history. All the proteins in question, protein 4.9 included, have close analogues in other tissues. Because the cytoskeletons of other cells are in general much more dynamic than the red-cell membrane skeleton, it appears altogether more likely that phosphorylation of their constituents may play some indispensable role in non-erythroid — neuronal, for example — contexts. □

1. Husain-Chishti, A., Levin, A. & Branton, D. *Nature* 334, 718–721 (1988).
2. Cianci, C.D., Giorgi, M. & Morrow, J.S. *J. cell. Biochem.* 37, 301–315 (1988).
3. Lu, P.-W. *et al.* *J. biol. Chem.* 260, 14958–14964 (1985).
4. Bennett, V. A. *Rev. Biochem.* 54, 273–304 (1985).
5. Mische, S.M. *et al.* *J. Cell Biol.* 105, 2837–2845 (1987).
6. Eder, P.S. *et al.* *Biochemistry* 25, 1764–1770 (1986).
7. Ferrell, J.E., Jr & Huestis, W.H. *J. Cell Biol.* 98, 1992–1998 (1984).

Lars Backman is in the Department of Biochemistry at the University of Umeå, Umeå, Sweden.

## Daedalus

# Seeing the strain

DESPITE great advances in the computer simulation of structures and finite-element analysis of components, engineering design is still a fairly black art. The actual distribution of stresses that a given load will induce in a structure, and the variety of loads that it will meet in service, can only be estimated. Every engineer offsets his ignorance with prudent but wasteful margins of safety.

Daedalus plans to change all this. DREADCO's chemists are synthesizing dyes whose chromophore groups, strung along a polymer backbone, are sterically very crowded. Even a very small dimensional change will cause a big alteration in the electronic overlap between neighbouring chromophore orbitals. As a result, the dye will change colour dramatically when squashed or stretched. The final product will be incorporated into DREADCO's strain-sensitive Loadlacker<sup>®</sup>.

Painted with Loadlacker and loaded, an engineering component will reveal its underlying stress pattern as a colourful map of contrasting hues, readily calibrated by DREADCO's shade card. It will blush coyly in regions of embarrassingly high stress but maintain a stern and unyielding colour where the metal is far too strong for the job. The designer will know immediately what modifications are needed.

Applied to a finished bridge or aircraft, say, Loadlacker will discount the static load due to the structure itself, and reveal only the additional service loads as they come on. A heavy lorry crossing the bridge will be accompanied by a subtle discoloration of the girders beneath it; the wings of the cruising aircraft will flush in a rainbow gradation from the highly stressed wing roots to the lightly loaded tips, and shimmer subtly with the buffeting of air-turbulence. In either case, excessive loading will signal its own warning; impending disaster will be averted.

Loadlacker may be even more popular in the domestic sector of the market. The rickety shelves and brackets put up by amateurs, and the precarious ladders they use to do so, could all benefit from its premonitory coloration. Loadlacker string, giving due warning that it is about to snap, should be especially popular. And fabric woven from Loadlacker thread could bring a whole new appealing dimension to the rag trade. Jogging outfits whose flashing and shimmering colours mirror their wearers' exertions; provocatively tight creations displaying a stress map of their tense equilibrium with the form inside; baby clothes whose gradually changing hue reveals when the baby has got too big for them — all will swell DREADCO's rosy bank account.

David Jones



## Another gas burst in a Cameroon lake?

SIR—The emission of gas, mostly CO<sub>2</sub>, from Lake Nyos on 21 August 1986 that killed about 1,700 people (Freeth, S.J. & Kay, R.L.F. *Nature* **325**, 104; 1987) drew attention to the large number of crater lakes in Cameroon, many of which have released gas. Since that incident was preceded by a gas burst from Lake Monoum on 15 August 1984, which killed 37 people, it is likely that many of the crater lakes have emitted gas bursts, or will do so.

In February 1987, we sampled sediment cores from the bottom of five lakes, respectively from south to north, Lakes Ossa, Barombi-Mbo, Bambuluwé, Oku and Nyos (Maley, J. *et al. J. Volcan. geotherm. Res.*; submitted). The measurement of the <sup>210</sup>Pb profile along each core allowed the sedimentation rate over the past century to be determined for three of these lakes, Ossa, Barombi-Mbo and Oku. The natural and artificial radioisotope profiles of the cores from the other two lakes contained certain anomalies (Piboule, M. *et al. J. Volcan. geotherm. Res.*; submitted).

In the case of Lake Nyos, we concluded that the sediment had been mixed for several metres in depth as there were constant values of <sup>137</sup>Cs and <sup>210</sup>Pb along the entire core length. There also appeared to be an excess of <sup>210</sup>Pb due to an additional flux of radon out of the sediment, as shown by a radioactive imbalance be-

tween <sup>210</sup>Pb and <sup>226</sup>Ra.

In Lake Bambuluwé, located about 10 km south of Bamenda, the <sup>210</sup>Pb measurements in sediments along the upper 15 cm exhibit a constant value of radioactivity (369 Bq kg<sup>-1</sup>). As in Lake Nyos, this activity is unusually strong. We have also observed gas bubbles on the core during its recovery from the bottom of the lake, indicating that its gas content was very close to the gas saturation at the sampling depth (58 m at Bambuluwé lake).

It therefore appears possible that Lake Bambuluwé is following the same pattern as Lake Nyos, that is an increase of radon flux through the lake sediments. Will this flux, as in Nyos, lead to a large CO<sub>2</sub> release? Our present measurements do not allow us to determine this with certainty. As well as completing these first determinations, it will be important to test by a new sampling if modification of the sediment column has taken place since 1987.

M. POURCHET  
J.F. PINGLOT

*Laboratoire de Glaciologie et  
Géophysique de l'Environnement,  
CNRS, Grenoble, France*

J. MALEY

*Laboratoire de Palynologie, ORSTOM,  
UR-A3, USTL, Montpellier, France*

M.A. MELIERES

*Laboratoire de Spectrométrie Physique,  
USTMG, Grenoble, France*

## Experimental design flaws still unaccounted for

SIR—For all the space and detailed analysis devoted to the 'Baltimore affair'<sup>1</sup>, the issue of the design of the experiments reported by Weaver *et al.*<sup>2</sup> has not been raised.

Suppose that in 1985 the possibility arose to pool the resources of two groups. One group had developed reagents that characterize the set of idiotype determinants expressed in immunoglobulin molecules formed in response to immune challenge with a hapten called NP. The other group had cloned a gene that codes for the completely assembled heavy chain of one of the set of idiotypically identifiable immunoglobulins formed in response to NP. The gene persisted in several copies in the descendants of at least two transgenic mice. The value of collaboration between the two groups to study immune regulation would have been evident. And so it must have seemed to the authors of the paper in question.

I assume that all these procedures gave absolutely clearcut results and that any unreported experiments were properly excluded because the experimenters were aware of trivial errors. Thus I ignore all of the issues discussed in ref. 1.

Presumably some hypothesis existed in the minds of the workers. All the collaborators may well have had different hypotheses. But some underlying question beyond the mindless "let us see what happens if you test our mice with your reagents" must have been framed. Given the state of knowledge of immune regulation in 1985 (or in 1988) any question raised or hypothesis stated required at least one specificity control. To make any sense at all out of immune manipulation of the transgenic mice, which carried an abnormally large number of copies of a pre-rearranged gene inserted in a variety of unnatural sites, required that comparisons be made with mice bearing a transgene of another (non-crossreacting) specificity. That would have been a minimum requirement for decent experimental design. Other mice expressing fewer copies of the transgene, for example, would have been desirable, but one must begin somewhere and there are practical considerations to take into account. Lacking the one basic control the discussion becomes pointless, and the arguments over technical issues a waste of time.

I believe the weakness in experimental

design, which is but one of many examples of the somewhat mindless application of the powerful techniques of molecular genetics and immunology, is more troublesome than the tempest in a teapot concerning the minutiae of test precision. With decent experimental design, the trivial questions would almost certainly not arise.

LEON T. ROSENBERG

*Department of Medical Microbiology,  
Stanford University School of Medicine,  
Stanford, California 94305-5404, USA*

1. Maddox, J. *Nature* **333**, 795-797 (1988).
2. Weaver, D. *et al. Cell* **45**, 247-259 (1986).

## Forced dilemma to come in muscle contraction

SIR—Malcolm Irving, in his interesting News and Views<sup>1</sup> comment on the results of Kishino and Yanagida<sup>2</sup>, does not point out a major dilemma that will result for the understanding of muscle contraction if their preliminary estimate of the force per crossbridge at less than 1 pN is confirmed.

Irving considers only the situation when the efficiency of energy conversion is in the region of 10%, but the experimental values for muscle are much higher than this, in the region of 50-60% for frog<sup>3,4</sup> and about 80% for tortoise<sup>5</sup>. The standard view of muscle contraction<sup>6</sup> is that each crossbridge cycle is driven by the splitting of one ATP. Thus, even using the lowest estimate of muscle efficiency, the work done by a single crossbridge in one cycle averages 50% of the free energy of splitting one ATP molecule, that is, at least  $43 \times 10^{-21}$  J. From measurements of the force changes produced by length steps<sup>7</sup> we know that the working stroke of a crossbridge is no more than 15 nm. It follows that the average force produced by a crossbridge during shortening must be at least 3 pN, and presumably during isometric contraction it will be higher.

If it turns out from direct measurements that the force per crossbridge is indeed less than 1 pN it will be necessary to suppose that one ATP molecule provides, on average, the energy for not one but several crossbridge cycles. This would require the myosin molecules to undergo a much more complex sequence of changes than anything now envisaged.

R. C. WOLEDGE

*Department of Physiology,  
University College,  
London WC1E 6BT, UK*

1. Irving, M. *Nature* **334**, 11 (1988).
2. Kishino, A. & Yanagida, T. *Nature* **328**, 74-76 (1988).
3. Kushmerick, M.J. & Davies, R.E. *Proc. R. Soc. B* **174**, 315-353 (1969).
4. Hill, A.V. *Proc. R. Soc. B* **126**, 136-195 (1938).
5. Woledge, R.C. *J. Physiol., Lond.* **197**, 685-707 (1968).
6. Huxley, A.F. *Reflections on Muscle* (Liverpool University Press, Liverpool, 1980).
7. Ford, L.E., Huxley, A.F. & Simmons, R.M. *J. Physiol., Lond.* **311**, 219-249 (1977).

## Relationships of humans to chimps and gorillas

SIR—The issues raised by Jared Diamond<sup>1</sup> concerning the phylogeny of humans and the apes turn on the interpretation of very little data. Given that humans, chimps, and gorillas are closely related, which pair are most closely related? The evidence linking chimps to gorillas subsumes some genetic data, anatomical data and most studies of the chromosomes. The only ambiguities — which Diamond takes as unequivocal proof of a chimp–human phylogenetic bond, as opposed to a chimp–gorilla bond — seem to exist in the interpretation of a very few highly publicized molecular studies.

I wish to raise the possibility that Diamond may have been misled by genetic braggadocio. Genetic studies have consistently failed to resolve the closest relatives among human, chimpanzee and gorilla, which has been a continuing source of frustration to practitioners. Given the general lack of humility common in molecular biological studies, the field's achievements may occasionally be difficult to distinguish from its aspirations.

The important problem in interpreting all of these results is to gauge which are giving us phylogenetic history, and which are not. Obviously only some are; but by selectively and uncritically reporting conclusions we stand to gain little.

Diamond's first line of evidence for the human–chimp bond is the  $\eta$ -globin pseudogene region. Sequencing a stretch of about 7,100 nucleotides in each of the four taxa (*Homo*, *Pan*, *Gorilla* and *Pongo*), Miyamoto *et al.*<sup>2</sup> located only eight with which to link humans and chimps. This was not many, but they properly recognized that it was more than the three they could find to link chimp and gorilla, or the three they found to link human and gorilla. The strength of these characters will have to outweigh two significant considerations. The first is homoplasy: there are, after all, only four nucleotides. The second is sequencing error, by which I do not impugn the technical competence of the authors, but rather note a widespread concern in the field<sup>3</sup>. Miyamoto *et al.*, themselves, observe in footnote 31 of their paper that in sequencing the human DNA, they found 44 sites at which their sequence differed from another determined sequence of the same clone. Where every nucleotide synapomorphy counts — as in this case — we need to be very circumspect about the basis upon which we rewrite our phylogenies. Superimpose upon this the spectre of nucleotide polymorphisms<sup>4</sup>, and I prefer the more conservative conclusion, that these sequence data are interesting, rather ambiguous (as are most of the other genetic data bearing on the problem) and have not felled the phylogenetic tree link-

ing chimps and gorillas.

The second line of evidence adduced by Diamond is the DNA hybridization study of Sibley and Ahlquist<sup>5,6</sup>. That study, unfortunately, is still not fully documented. In my experience the melting profiles, per cent hybridizations, melting temperatures and even the analytical procedures on which Sibley and Ahlquist based their conclusions are not available to interested colleagues. Sibley and Ahlquist, since first publicizing their conclusions, have consistently failed to meet the burden of proof which falls upon all investigators. Until that burden is met, it is gratuitous to assume the interpretations are valid, or to draw conclusions from it.

Invoking Sibley and Ahlquist's conclusions, however, Diamond then proceeds from the proposition that humans are related to chimpanzees, to infer that humans actually are chimpanzees. Chimpanzees obviously have an intimate relationship of common ancestry with humans (and gorillas, which have been occasionally grouped as congeneric with chimps) but common ancestry is not identity. I may be my brother's keeper, but I am not my brother.

This absurd declaration is not the first such statement by a molecular advocate attempting to generalize the genetic similarity of humans, chimps and gorillas. In fact, Zuckerkandl<sup>7</sup> deduced a quarter century ago that "from the point of view of haemoglobin structure, it appears that gorilla is just an abnormal human, or man an abnormal gorilla, and the two species form one continuous population".

Yet we are not our haemoglobin. Scientific progress entails distinguishing between hyperbole and description; between the literary and the literal — and not confusing one for the other. Is it possible that the molecular approach has not brought us as far in the past twenty-five years as is commonly thought?

JON MARKS

*Departments of Anthropology  
and Biology,  
Yale University,  
New Haven,  
Connecticut 06520, USA*

1. Diamond, J. *Nature* **332**, 685–686 (1988).
2. Miyamoto, M. *et al. Science* **238**, 369–373 (1987).
3. Orkin, S.H. *Trends Genet.* **3**, 207 (1987).
4. Savatier, J. *et al. J. molec. Evol.* **24**, 309–318 (1987).
5. Sibley, C. & Ahlquist, J. *J. molec. Evol.* **20**, 2–15 (1984).
6. Sibley, C. & Ahlquist, J. *J. molec. Evol.* **26**, 99–121 (1987).
7. Zuckerkandl, E. in *Classification and Human Evolution* (ed. Washburn, S.L.) (Wenner-Gren, New York, 1963).

DIAMOND REPLIES—I discussed<sup>1</sup> two issues: which pair among human, chimp and gorilla are each other's closest relatives; and the relative merits of DNA sequencing and DNA/DNA hybridization for deducing phylogenies. I agree with the reservations of Marks (and Filipowski<sup>2</sup>) about DNA sequencing, and it was on the basis of some of the same points that I concluded by preferring DNA/DNA

hybridization of the whole genome over sequencing of genome fragments as a phylogenetic method. But I disagree with Marks' comments about hominoid relationships.

Marks begins by stating that the only ambiguities in the weight of evidence linking chimps to gorillas are a very few molecular studies. In fact, of course, the non-molecular evidence is equivocal and diverse lines of molecular evidence show that human, chimp and gorilla are close but that human and chimp are slightly closer to each other than either is to gorilla. To previous molecular evidence<sup>3</sup> based on mitochondrial DNA, DNA/DNA hybridization,  $\eta$ -globin pseudogenes, fibroblast polypeptides and immunoglobulin E pseudogenes, can be added recent data from a further eight kilobases of sequence<sup>4,5</sup>.

While Marks refers to the study of Sibley and Ahlquist as undocumented, these authors gave detailed descriptions of their methods in many earlier papers and presented their hominoid data in two lengthy papers, of which ref. 3 gives 514 DNA/DNA hybridization values. At Marks' urging, J. Powell and A. Caccone at a recent meeting of the Society for the Study of Evolution redetermined hominoid DNA/DNA hybridization values by a different method, using some samples provided by Sibley and Ahlquist as well as others obtained with Marks' help, and obtained results concordant with those of Sibley and Ahlquist<sup>6</sup>.

Finally, Marks attributes to me the inference that humans are chimpanzees, which he labels an absurd declaration. Instead, I noted that common chimps, pygmy chimps and humans are each other's closest relatives, and that the DNA differences among these three species are no greater than the differences among species assigned to the same genus in other animal groups<sup>6</sup>. Congeneric does not mean identical.

JARED M. DIAMOND

*Physiology Department,  
UCLA Medical School, Los Angeles,  
California 90024, USA*

1. Diamond, J. *Nature* **332**, 685–686 (1988).
2. Filipowski, J. *Nature* **334**, 571 (1988).
3. Sibley, C.G. & Ahlquist, J.E. *J. molec. Evol.* **26**, 99–121 (1987).
4. Maeda, N. *et al. Molec. biol. Evol.* **5**, 1–20 (1988).
5. Miyamoto, M. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
6. Diamond, J. *Nature* **310**, 544 (1984).

### Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of less than 500 words and five references. □



# Why do we do as we do?

Philip Kitcher

**Patterns, Thinking, and Cognition: A Theory of Judgment.** By Howard Margolis. University of Chicago Press: 1988. Pp.332. Hbk \$45, £35.95; pbk \$15.95, £12.95.

ONCE there was formal logic, an instrument of shining beauty. Developed at the turn of the century by Gottlob Frege and Bertrand Russell, it was designed primarily to clarify mathematical reasoning and was apparently successful in doing so. Then came probabilistic inference, which was somewhat less elegant but offered models of rational decision-making that might prove applicable across a range of scientific and everyday contexts. So, in the late twentieth century, we can hope to fulfil a leibnizian dream, to see thinking as a species of symbol-manipulation and to resolve disputes by calculation. There are areas in which improvements are needed — our understanding of inductive and analogical reasoning is imperfect — but, given time, statisticians, philosophers and the artificial intelligentsia will remove any remaining warts from our image of human reasoning and cognition.

Howard Margolis believes that the project is wrong-headed and our confidence in its successes misplaced. He has written a stimulating book that attempts to make us see things differently.

Margolis begins with the idea that thinking and judgment are species of pattern-recognition. So, what is a pattern and what is pattern-recognition? We are not told, for, as Margolis admits, "no one can say much yet about what the brain is doing when it recognizes a pattern" (p. 3). But his reluctance to speculate does not doom the enterprise, for the central task of the first part of the book is to relate the primitive concept of pattern-recognition to other psychological notions and to aspects of human behaviour. The initial step is to view the activation of patterns as primary, the provision of reasons being a subsequent way of rationalizing judgment either for ourselves or for others. In Margolis's telling analogy, our sincere accounts of our reasons for arriving at a particular decision are like the protestations of the post-hypnotic patients who find excellent excuses for jumping up and scratching their ears. But, in our case, it is the cueing of a pattern, not the suggestions of the hypnotist, that causes the judgment or decision.

Pattern-recognition, Margolis claims, occurs among most multi-celled animals and has an evolutionary history of over 500 million years. An important project for the early part of the book is to explain how various types of cognitive process — learning, choice, judgment, reasoning and

calculation — can be understood as successive developments of pattern-recognition. Thus Margolis attempts to describe a 'cognitive ladder' that will lead from the simplest types of pattern-recognition, which he sees as being distributed widely across the animal kingdom, up to the more specialized types that occur (so far as we know) only in our own species. The description is self-consciously darwinian. Margolis emphasizes the continuity of the processes in his sequence, and endeavours to relate the innovations to hypothetical adaptive advantages.

We can now make more explicit the contrast between intuitive judgments and the 'logical justifications' that people offer for them. On Margolis's account, an intuitive judgment occurs when a pattern is activated: something comes to 'look right' to us. We may reflect on the situation, offering reasons for judging as we do, and then examine those reasons. At its most explicit, the process may involve appeal to the canons of formal logic, an appeal that Margolis construes as the activation of recondite patterns, cued by linguistic similarities, that can have emerged only late in the history of human pattern-

recognition. The reflections may prove congruent with the original judgment, reinforcing our sense that we have judged correctly, or they may be at odds with it. Using the potential for match and mismatch, Margolis offers an interesting account of many vexed cognitive concepts (belief, knowledge, doubt and so forth) and an analysis of learning.

Although his presentation is illustrated by familiar examples from the psychological literature, it is only in the second half of the book that Margolis begins systematically to assemble evidence for his claims. One source of support is a re-evaluation of the famous tests designed by Wason, Johnson-Laird, Tversky, Kahnemann and others to expose foibles of human rationality. Margolis brings a genuinely novel approach to this material. Unlike many other writers, he does not trace our cognitive lapses to the abstract formulations in which some of the tasks, such as Wason's card-selection experiment, are presented — there is no deep problem in acquiring principles of formal reasoning. Margolis believes that people can come to employ those sophisticated patterns that are activated in the doing of logic.

The troubles stem from the fact that we have other patterns at our disposal which are cued when we are set certain types of problem. The consequence is that we sometimes fail to reach the conclusions that logicians recommend. Margolis suggests that two separate kinds of factor promote the miscueing: *semantic* factors



*Flights of fancy* — in 1835 the New York Sun, a small local newspaper, published a series entitled "Remarkable discoveries made at the Cape of Good Hope". Claiming to be extracts from articles by Sir John Herschel in the Edinburgh Journal of Science (not then in publication), the pictures showed in detail the 'batmen' inhabiting the Moon. The paper's sales shot up until the hoax was exposed as the work of the journalist Richard Adams Locke, but not before being reprinted and translated into several languages. The picture is taken from Under Capricorn: A History of Southern Hemisphere Astronomy by David S. Evans, published by Adam Hilger, price £29.50; in the United States distributed by Taylor & Francis, \$78.

that allow us to interpret the words used in posing the problem in an unintended fashion; and *scenario* effects that incline us (mistakenly) to assimilate the task to quite a different project than any standard interpretation of the words would permit. He describes modifications of the original experiments that would be expected to generate much higher success rates (modifications that block cueing of the wrong patterns) and reports some of the results.

The last six chapters deal with a different type of evidence. Margolis looks to the history of science, in particular to the great revolutionary changes in science, to see if the formulation, defence and discussion of radically new ideas can be illuminated by his account of cognition. Here, success depends on showing that those who originate new concepts (Copernicus and Darwin are the exemplars) were in an especially good position to do so, on showing how their arguments induced the right kinds of patterns in those who supported the new cause, and on explaining the resistance of the opposition.

Margolis devotes most attention to the copernican revolution. He argues that Copernicus's route to discovery could indeed have begun with his professed dissatisfaction with the use of equants in Ptolemaic astronomy, even though the equant is logically irrelevant. Copernicus was particularly well-placed to articulate a new pattern (a view of the Solar System in which the ordering of the planets was both determinate and consonant with the periods of rotation) because of his freedom from teaching, because of his own early research on the anomalies of the motions of the Moon and because of the impact of new discoveries in geography. Together, Margolis says, these factors would have diminished Copernicus's commitment to the ptolemaic pattern of viewing the Solar System as a group of nested spheres. Once freed from that commitment, he was able to develop an alternative perspective until its intrinsic harmony became apparent. Moving the Earth is a small by-product of the new vision.

Margolis continues with some thought-provoking analyses of the reception of copernicanism and the trial of Galileo. There are places in which his history seems strained or lop-sided: in his neglect of the standard explanation (Kepler's predilection for bizarre numerology) for Galileo's failure to refer to Kepler's laws (p. 286), and especially in his downplaying of the role of aristotelianism in the copernican debate. But these are minor blemishes on a provocative reinterpretation.

More fundamental is a difficulty which besets not only the historical analysis but much of the book. It is often hard to know whether the claims that are made about

subjects' performances, in twentieth-century academia or Renaissance *palazzi*, really involve the cueing and miscueing of patterns. At various key stages in the book Margolis describes behaviour that could more easily be analysed in everyday terms or in the idioms of other cognitive theories. Does the notion of pattern-recognition appear so powerful because it is so plastic? Could a clever exponent deploy it to cover anything?

Naive falsificationists, posing such questions, might conclude that this book is

worthless. They would be wrong. *Patterns, Thinking, and Cognition* is complex and compressed, but the scope and suggestiveness of Margolis's proposals indicate that the effort involved in making them more precise and definite may be very worthwhile indeed. We are offered a broad and challenging vision of human cognition. Formal logic and its offspring may never look the same again. □

*Philip Kitcher is a Professor in the Department of Philosophy, University of California at San Diego, La Jolla, California 92093, USA.*

## Ways of knowledge

Drew McDermott

**Logical Foundations of Artificial Intelligence.** By Michael R. Genesereth and Nils J. Nilsson. *Morgan Kaufmann: 1987. Pp.405. Distributed by W.H. Freeman, \$36.95, £29.95.*

ANYONE interested in artificial intelligence (AI) should be interested in the topic of this book, that is, computer manipulation of knowledge expressed in logical formalism. Unfortunately, the book is at best a partial success.

Knowledge is important to AI because an intelligent agent must know something about the world it lives in, and about how to solve problems that arise in that world. Many practitioners of AI, including the authors of this book, are driven by the idea that knowledge should be represented explicitly — 'declaratively'. And many who espouse that position further believe that mathematical logic is the proper framework for studying the syntax, semantics and inference mechanisms of declarative knowledge representation systems. This is the position known as 'logicism'. Genesereth and Nilsson place themselves in that camp.

It would be good to have a thorough exegesis of logicism, which enjoys a position of prestige as one of the few theoretically weighty branches of AI. But the book doesn't fill that role. It fails to present an adequate case for why AI should concentrate so strongly on deduction as the paradigm for inference. Worse, it is rather parochial and fails to do justice to a lot of logicist research. Many people outside Stanford will be surprised to find that their work is either not mentioned at all or is relegated to bibliographical notes.

The book does have some high points. Chapter 1 is a concise presentation of the logicist position, and in Chapter 2 Genesereth and Nilsson do a good job of explaining denotational semantics for AI purposes. The notion of *conceptualization*, which the authors depend on, is a nice adaptation of the idea of *intended model* to AI. But they are not sure what to

say about inference. They classify inference processes into two groups, sound and unsound, and acknowledge that most interesting inference algorithms are unsound (liable to draw invalid conclusions), but their attention is focused on the sound ones such as resolution theorem proving. So, after the presentation of deductive techniques in Chapters 3 to 5, the discussion bifurcates into treatments of various nondeductive forms of inference, and unrelated applications of deduction to particular domains. As an example, after the illuminating discussion of 'nonmonotonic' logic in Chapter 6 (that is, logic that allows defeasible conclusions to be drawn), no use is made of it again. Instead, in Chapters 11 and 12, on temporal reasoning and planning, problems are defined in terms of classical monotonic deduction where possible. For instance, the planning problem is defined as the problem of proving that a series of actions will bring about a desired state of the world. This is not a definition that an active researcher in the field would accept.

There are a few places where the presentation is lucid and thorough. The account of nonmonotonic logic in Chapter 6 is instructive, especially the parts about John McCarthy's 'circumscription' technique. This chapter is as up to the minute as a textbook can hope to be. Chapter 9, on the logic of knowledge and belief, is another winner. Here, two approaches are described: Kurt Konolige's syntactic approach, in which belief is a predicate on formulae in an agent's database; and Jaakko Hintikka's possible-worlds approach, adapted to AI by Robert Moore, in which belief in a proposition is attributed to an agent if the proposition is true in all possible worlds compatible with what he believes. Both 'conceptualizations' are interesting and powerful.

Sadly, however, many of the other chapters are either out of date or aimless. The reader is not given much internal evidence of the value of the logicist approach, and may find his doubts about it unassuaged. □

*Drew McDermott is a Professor in the Department of Computer Science, Yale University, PO Box 2158, Yale Station, New Haven, Connecticut 06520, USA.*



# In the business of synthesis

Robert M. Joyce

**Petrochemicals: The Rise of an Industry.** By Peter H. Spitz. Wiley: 1988. Pp. 588. \$29.95, £25.95.

OF THE many changes in the American way of life triggered by the Japanese attack on Pearl Harbor in 1941, little publicity has been given to one of the outstanding technological achievements of the modern era. Within two years, engineers had designed, constructed and put into operation plants that increased US production of 100-octane aviation gasoline and the new synthetic rubbers, GRS, Buna S and neoprene, by factors ranging from 15 to 180 over prewar levels. Concurrently, plants had to be built to produce the intermediates on which these materials were based: lower olefins for alkylate, and styrene, butadiene, isobutylene and acetylene for the rubbers.

*Petrochemicals* is primarily the story of the dramatic expansion of the US chemical industry after 1941. As Spitz makes clear, the experience and shared information gained by the companies involved in the wartime feat were key factors in the subsequent explosive growth of the industry.

The book begins with historical accounts of the origin, primarily in Germany, of the synthetic chemical industry based on coal tar intermediates, and of its prewar status in the United States. The latter history is framed in vignettes of four companies which were prime movers of the US industry into petroleum raw materials: Union Carbide, Shell Chemical, Dow and Standard Oil of New Jersey. Some of the history is inaccurate, and is flawed by technical errors. On tetraethyl lead, Spitz states (p. 91): "General Motors and Standard Oil (N.J.) subsequently founded the Ethyl Corporation and a new industry was born". Actually, General Motors first came to Du Pont, who developed a process for TEL and commercialized it in 1923; Ethyl Corporation was essentially only a legal entity until construction of their first plant in 1936. And (to give just one example) Cellosolve (ethoxyethanol) is identified as an ether based on ethylene oxide and cellulose (p.79).

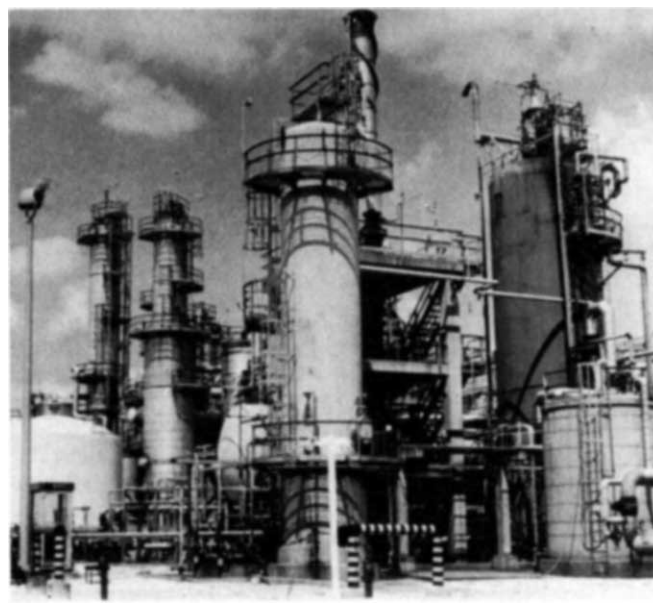
The next two chapters describe the developments of the late 1930s and early 1940s — catalytic cracking, platforming and separations technology — that put petroleum feedstock on a firm footing as a source of chemical intermediates. These accounts are salted with interesting biographical portraits of seminal contributors to the fields, and illustrated with many complex unit-process diagrams. As

a necessary prelude to discussing the postwar industry, Spitz intercalates an interesting chapter on the cartels that characterized much of the industry in the late nineteenth and early twentieth centuries. By the 1930s, however, cartels had come into disfavour in the United States, and by the end of the Second World War the stage was set for all comers to seek the perceived gold to be found in the new synthetic chemicals and polymers.

Spitz provides some historical background and descriptions of the underlying technical advances. Here again there are errors. Thus (p. 279): "Paul Schlack — of IG — started work on polycondensation in 1928, several years before Carothers"; actually, Carothers came to Du Pont in

competition, not only from other US firms but also from aggressive companies in Europe and Japan and, later, from oil-producing countries in the Middle East. Spitz's description of the rise and shakeout of the vinyl chloride industry is an object lesson for planners and managers.

The best chapter in the book is the last, in which the author describes how the industry seems finally to have assimilated the lessons of the 1950s–1970s, in no small measure the hard way. Many older, uneconomical petrochemicals plants have been shut down, and a large number of petrochemicals manufacturers have abandoned or sold their commodity chemicals businesses. His analysis of the



*Engineering meets chemistry — part of Polysar's ethylbenzene unit in Sarnia, Ontario. The picture is taken from Industrial Aromatic Chemistry by H.-G. Franck and J. W. Stadelhofer, just published by Springer-Verlag. Price is DM128, £44.*

February 1928 and began work on polycondensation almost immediately.

Synthetic fibres and plastics are identified as the triggers for the postwar expansion of the chemical industry. The boom was fuelled in part by public perception that nylon, polyethylene and synthetic rubbers were not merely 'ersatz,' but were useful materials with unusual properties. In addition, intermediates based on petrochemicals were required to make them, and many companies, particularly those in the petroleum business, quickly came to view synthetic chemicals and polymers as opportunities to broaden their base. Much key technology could be gleaned from government reports on the postwar inspection of German plants. The only patent protection in the area was on proprietary catalysts and process details, and these patents could often be circumvented by hiring an engineer from a competitor or by drawing on the knowledge of a contract firm that had built a competitor's plant.

Many US firms who joined the 'gold rush' into petrochemicals found it at best marginally profitable because of heavy

problems — some self-inflicted — in the growth of the industry and of their underlying causes is instructive reading for practising engineers and executives.

But this is not a satisfying book. It suffers from poor editing and proof-reading — many chemical names are improperly hyphenated or improperly separated into two more words, different names are used for the same compound and isomer indicators are not set in italics. The chemical index appears to have been created by a computer search of the text without any judgment being brought to bear on the result. Thus (and there are many more such examples) "metallic bromine" (!) from p. 327 of the text is listed under "M".

Altogether, the author would have done well to heed his own words (p. 132) about his teacher W. K. Lewis: "He taught his students to be clear and concise in their communications as engineers, particularly when addressing non-technical people". □

Robert M. Joyce, 1518 Valley Forge Boulevard, Sun City Center, Florida 33570, USA, is editorial coordinator of the book series Organic Reactions and a former director of research for Du Pont.

# Talking man to man

Roy Porter

**Homosexuality: A Philosophical Inquiry.** By Michael Ruse. Basil Blackwell: 1988. Pp. 292. £19.50, \$19.95.

IT WOULD be easy to be dismissive about a philosopher — such as Professor Ruse — who presumes to elucidate the problems of homosexuality through the light of reason. Does the subject fall within the domain of philosophy at all? Isn't it rather (it might be objected) a matter either of endocrinology or of inclination? Yet these are doubts which Professor Ruse largely scotches. For one thing, the scientific data remain so scattered and ambiguous that they require particularly cool evaluation. For another, the more that positions polarize over gay rights in this age of AIDS, the more those who shape opinion and frame legislation are duty-bound to transcend rant and cant. All of us at least should think straight.

So what does philosophy tell us? Ruse is sceptical of those radical critics who contend that 'homosexuality' is best viewed not as a natural fact but rather as a stigmatizing label. Same-sex erotic orientation is a basic reality of human societies, to say nothing of the animal kingdom; hence to identify people as 'homosexuals' is not to collude in some nasty plot. How we should explain homoeroticism is, however, altogether less clear.

Ruse surveys the field. He finds some plausibility in psychodynamic models (dominant mother, weak father): *pace* Popper, Freudian theory should be testable, Ruse argues, and such surveys as do exist lend it a degree of support. But psychosocial hypotheses take you only so far. The endocrinological evidence unambiguously points, Ruse concludes, to a hormonal component as well. Less convincing and coherent have been attempts to explain homosexuality in terms of sociobiological survival strategies.

Such aetiological matters obviously weigh heavily when Ruse takes up the hoary question as to whether homosexuality should be viewed as an illness — or even as a disease (the American Psychiatric Association defined it as a mental disorder up to 1974, when it reversed its view by postal ballot). While, surely rightly, making short shrift of the 'disease' argument, Ruse dwells more upon the issue of 'sickness,' noting how surveys show that homosexuals seem less happy, more suicide-prone, than heterosexuals. But here Ruse does not, to my satisfaction, explain how we could tell whether the cause of this distress is homosexuality

*per se*, or rather living under the threat of a ('sick') homophobic society.

Is there, then, anything 'wrong' with homosexuality? Here Ruse is basically liberal. Ethically speaking, both Kantian arguments (the sovereignty of the moral self) and utilitarianism (every person is the best judge of his or her own happiness) require that homosexuals must enjoy equal rights with heterosexuals. Yet he also has some hesitations.

Often evoking the image of the AIDS-prone bath-house gay, with his thousands of casual partners, Ruse implies that there may be something ultimately tawdry and inadequate about the homoerotic lifestyle. Yet here he is guilty of sleight of hand, shifting the argument from the ethics of homosexuality to the problems of promiscuity. Why should we presuppose

any intrinsic connections between the two? After all, as Ruse himself notes, lesbians tend to be less promiscuous than heterosexual males. Indeed, in these days of AIDS, lesbianism seems the healthiest sexual orientation of all. (In its relative indifference to lesbianism, the book shows a deplorable sexist bias.)

Two cheers, then, for this survey. It may do some good by showing that informed discussion is possible on a prickly topic. But I do hope that Professor Ruse's next book will be called *Heterosexuality*, just to prove that he is even-handed enough to recognize that male-female sexuality poses equally daunting conceptual and ethical problems. □

Roy Porter is at The Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BP, UK.

## In formation

H. Frederik Nijhout

**Insect Morphogenesis.** By Fritz E. Schwalm. Karger: 1988. Pp. 356. £131.90, \$193.50, SwFr. 290, DM 347.

THE introduction and back cover of this work declare that it was stimulated in large measure by the recent dramatic progress in the developmental genetics of *Drosophila*. Contrary to the expectations that such statements might raise, the book is not an attempt to integrate the recent work on *Drosophila* with the broader field of insect embryology. Rather, Dr Schwalm's purpose is to provide an overview of insect embryology from a strictly morphological viewpoint and to attempt to distil major themes (or, as he insists, the 'typical' form of the event) from the enormous diversity of developmental processes and adaptations that have evolved among the members of this vast taxon.

The book is divided into two parts of equal length. The first deals with adult morphology, gametogenesis and reproductive biology, while the second is a review of comparative insect embryology. The exact purpose of the first part within a book on insect morphogenesis is unclear. It is too long to be regarded as background for insect development but too short to allow adequate discussion of the topics included. For instance, juvenile hormone is mentioned as being involved in egg maturation, but little is said about the many different and wonderful control pathways for egg development that have now been elucidated, nor can the reader gain access to that literature through the references provided. Likewise, a section on sex determination mentions a few chromosomal mechanisms but ignores the extensive work on *Dros-*

*ophila* and the intriguing older material of Goldschmidt that is truly begging for a modern review.

The author has unfortunately passed up many exciting opportunities for synthesis in this part of the book. The curiously haphazard organization is probably best exemplified by noting that the section entitled "Nucleic Acid Synthesis and Storage During Oogenesis" contains substantial discussions of endosymbiosis, paedogenesis and chorion formation.

The second half of the book is a thorough account of comparative insect embryology from cleavage to hatching. The emphasis is on comparisons of embryonic structures and events between representatives of many insect orders. The author's attempts to draw generalities or themes out of comparisons of published work, however, more often than not stumble on the fact that equivalent observations have seldom been made on a sufficient diversity of insects to allow us to interpret with certainty what is primitive and what is derived, at what point in insect evolution a particular phenomenon first appeared and how generalized it is among the presumptive descendant taxa. Thus, the comparative approach, rather than enlightening us about the origin of various developmental phenomena, serves to focus our attention on the many gaps in our knowledge of their diversity and prevalence.

Throughout the book the author provides extensive tabulations of the manifestation of various developmental phenomena among the many taxa of insects. Such tables provide helpful summaries and constitute a rudimentary database that will be of use to anyone interested in comparative aspects of insect embryology. □

H. Frederik Nijhout is a Professor in the Department of Zoology, Duke University, Durham, North Carolina 27706, USA.



# The molecular heterogeneity of protein kinase C and its implications for cellular regulation

Yasutomi Nishizuka

Department of Biochemistry, Kobe University School of Medicine, Kobe 650, Japan

*Protein kinase C is now known to be a large family of proteins, with multiple subspecies that have subtle individual enzymological characteristics. Some members of the family exhibit distinct patterns of tissue expression and intracellular localization; different kinases probably have distinct functions in the processing and modulation of a variety of physiological and pathological responses to external signals.*

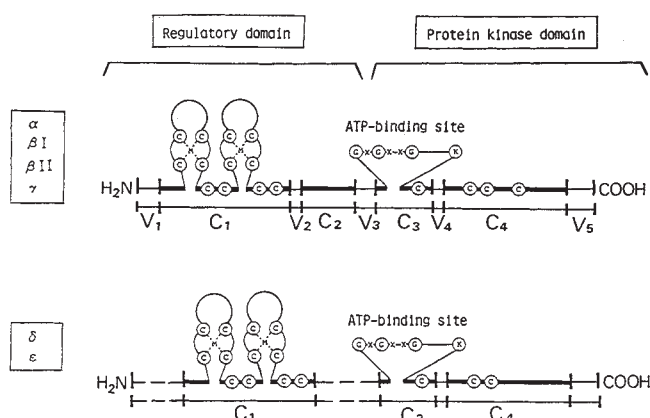
SINCE its discovery in 1977, the enzyme protein kinase C (PKC) has become a focus of attention for biologists interested in the mechanisms by which external signals are transmitted to the interior of cells, and in tumorigenesis. When a ligand binds to certain receptors on the cell surface, inositol phospholipids are hydrolysed, producing diacylglycerol and inositol phosphates, both thought to act as 'second messengers'. It is thought that the primary effect of the diacylglycerol is to activate PKC which in turn phosphorylates a range of cellular proteins. It is especially interesting that the enzyme also serves as the receptor for phorbol esters, a class of tumour promoter; it is probably the inappropriate signals given through PKC which result in the tumour-promoting characteristics of these compounds.

It is now clear that there is more than one species of PKC molecule, and several discrete subspecies have recently been defined. There are probably more to discover, and we are only just beginning to understand how the individual subspecies may have different functions. Here I briefly review what is known so far, in the hope of encouraging the resolution of some of the outstanding problems.

## Members of the family

At the last count, seven subspecies of PKC have been identified (Table 1). The integrated nomenclature I use herein for the PKC subspecies known at present is established elsewhere<sup>2-4</sup>. Four subspecies ( $\alpha$ ,  $\beta_1$ ,  $\beta_{11}$  and  $\gamma$ ) emerged from the initial screening of a variety of complementary DNA libraries<sup>5-13</sup>. More recently, at least three further subspecies ( $\delta$ ,  $\epsilon$  and  $\zeta$ ) have been isolated from a rat brain library by using a mixture of  $\alpha$ ,  $\beta_{11}$ , and  $\gamma$  complementary DNAs as probes under low stringency conditions<sup>4,14</sup>. (An  $\epsilon$ -cDNA clone has also been isolated from a rabbit brain library<sup>15</sup>.) These subspecies have a common structure closely related to, but clearly distinct from, the four subspecies initially described. The structures of the two groups are shown schematically in Fig. 1.

The conserved region C<sub>1</sub> contains a tandem repeat of a cysteine-rich sequence, Cys-X<sub>2</sub>-Cys-X<sub>13(14)</sub>-Cys-X<sub>2</sub>-Cys-X<sub>7</sub>-Cys-X<sub>7</sub>-Cys, where X represents any amino acid. (The  $\zeta$ -subspecies has only one set of the cysteine-rich sequence; Y. Ono *et al.*, manuscript in preparation). This sequence is similar to the consensus sequence of a cysteine-zinc-DNA-binding finger that is found in many metallo-proteins and DNA-binding proteins that are related to transcriptional regulation<sup>16</sup>, but there is no evidence that PKC binds to DNA. The carboxy-terminal half of each enzyme containing the regions C<sub>3</sub> and C<sub>4</sub>, seems to be the catalytic domain, as it shows large clusters of sequences that resemble many other protein kinases<sup>4,5,7-15,17</sup>. The conserved region C<sub>3</sub> has an ATP-binding sequence, Gly-X-Gly-X-X-Gly...Lys. Although the conserved region C<sub>4</sub> contains a similar sequence, Gly-X-Gly-X-X-Gly...(X), its significance is unknown. The sites involved in binding to calcium, diacylglycerol and phospholipid are presumably contained in C<sub>1</sub> and C<sub>2</sub> but have not yet been identified with any certainty.



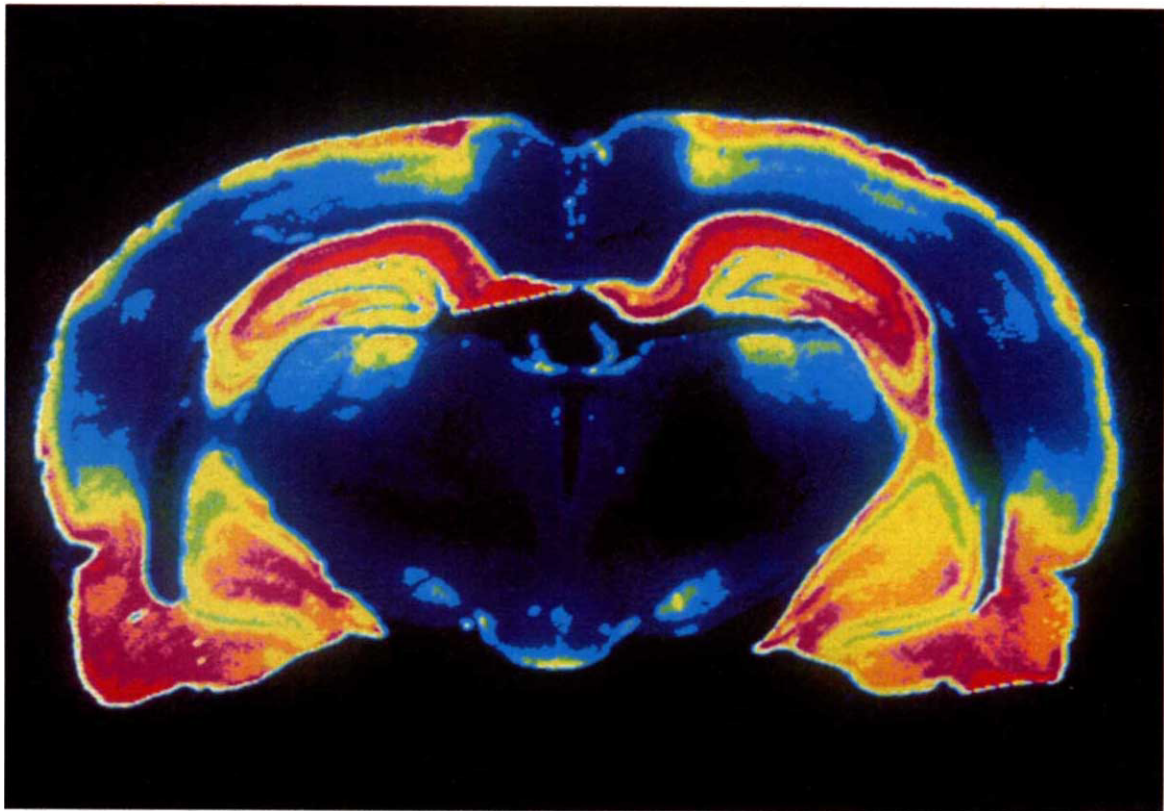
**Fig. 1** Common structure of PKC subspecies. C, G, K, X, and M represent cysteine, glycine, lysine, any amino acid, and metal, respectively. Four conserved (C<sub>1</sub>-C<sub>4</sub>) and five variable (V<sub>1</sub>-V<sub>5</sub>) regions of the larger group are indicated. The main difference between the two groups is that the  $\delta$ ,  $\epsilon$  and  $\zeta$  subspecies lack the second conserved region C<sub>2</sub>. The  $\beta_1$  and  $\beta_{11}$  subspecies, which seem to be derived from a single messenger RNA by alternative splicing, differ from each other only in ~50 amino acid residues at their carboxy-terminal end regions, V<sub>5</sub>, and even in this area they possess a high degree of sequence homology<sup>2,7-11</sup>.

PKC was originally detected as an undefined protein kinase, which could be activated by limited proteolysis with the Ca<sup>2+</sup>-dependent neutral protease, calpain<sup>18</sup>. This hydrolytic cleavage occurs at one or two specific sites in the region V<sub>3</sub>, resulting in the release of a fully catalytically active fragment, which is rapidly removed from the cell. The physiological significance of this proteolysis has not been unequivocally established but, for reasons discussed later, it may be related to down-regulation of the PKC molecule itself.

## Distinct expression and localization

Early studies of tissue distribution using tritiated phorbol-12,13-dibutyrate as a probe for PKC revealed an uneven distribution of PKC in the brain<sup>19,20</sup>. Northern blot analysis with specific oligonucleotide probes has suggested that some PKC subspecies are expressed specifically in certain tissues<sup>6,7,10-12</sup>, as has *in situ* messenger RNA hybridization<sup>21</sup>. Using biochemical, immunological and cytochemical procedures with subspecies-specific antibodies, the relative activity and individual pattern of expression of multiple PKC subspecies in several tissues have recently been extensively examined in this and other laboratories<sup>22-34,35</sup>. Much is known about the distributions of  $\alpha$ ,  $\beta$  and  $\gamma$ ; rather less about the  $\delta$ ,  $\epsilon$  and  $\zeta$ -subspecies.

Perhaps the most striking result is that in the rat, the  $\gamma$ -subspecies appears to be expressed solely in the brain and spinal cord, and is particularly concentrated in the hippocampus, cere-



**Fig. 2** Colour image analysis of immunoreactive material stained with a  $\gamma$  subspecies-specific antibody in a frontal section of the rat mid-brain. Red, region most heavily stained with this antibody. The detailed experimental conditions are similar to those described elsewhere<sup>28</sup> (photograph from ref. 28).

bral cortex and amygdaloid complex (Fig. 2). An almost identical distribution pattern is found in the monkey brain<sup>25</sup>, but here the cerebellar cortex (Purkinje cell bodies, dendrites and axons, see Fig. 3a) also contains stains heavily<sup>20-28,35</sup>. More specifically, the  $\gamma$ -subspecies seems to be localized in the nerve endings of the Purkinje cell axon that terminate in the deep cerebellar nuclei, but not in the post-synaptic cell body itself<sup>29</sup>. It is associated with membranous structures of the cell in general, except for the nucleus and perhaps mitochondria, which express little PKC. In the brain, the quantity of  $\gamma$ -subspecies increases postnatally, reaching a maximum in the rat at around three weeks after birth<sup>30</sup>. Although it seems premature to discuss the function of this PKC subspecies, it may be important in special-

ized neuronal processes, such as long-term potentiation in the hippocampus.

The tissue distribution of the  $\beta_I$  and  $\beta_{II}$  subspecies shows that there is much more  $\beta_{II}$  than  $\beta_I$  in the brain and many other tissues, including endocrine tissues such as pancreatic islets and the pituitary gland. Cytochemical analysis with polyclonal antibodies that distinguish the two subspecies reveals a distinct pattern of cellular expression in certain tissues<sup>31</sup> (see also refs 24, 26). For example, in the rat cerebellar cortex, the  $\beta_I$  subspecies is present mainly in the granule cell body (Fig. 3b) and the  $\beta_{II}$  subspecies primarily in the molecular layer, where it is apparently localized in the presynaptic nerve endings terminating on the dendrites and cell body of the Purkinje cell (Fig. 3c).

**Table 1** Subspecies of protein kinase C from mammalian tissues

| Subspecies                     | $\alpha$                                      | $\beta_I$              | $\beta_{II}$           | $\gamma$                     | $\delta$                  | $\epsilon$                | $\zeta^*$                 |
|--------------------------------|-----------------------------------------------|------------------------|------------------------|------------------------------|---------------------------|---------------------------|---------------------------|
| Amino-acid residues            | 672                                           | 671                    | 673                    | 697                          | 673                       | 737                       | 592                       |
| Calculated molecular weight    | 76,799                                        | 76,790                 | 76,933                 | 78,366                       | 77,517                    | 83,474                    | 67,740                    |
| Chromatographic sub-fraction   | type III                                      | type II                |                        | type I                       | not identified            | not identified            | not identified            |
| Chromosome location (human)    | 17                                            | 16                     |                        | 19                           | ?                         | ?                         | ?                         |
| Activators†                    | PS+DG+Ca <sup>2+</sup><br>AA+Ca <sup>2+</sup> | PS+DG+Ca <sup>2+</sup> |                        | PS+DG+Ca <sup>2+</sup><br>AA | PS+DG+(Ca <sup>2+</sup> ) | PS+DG+(Ca <sup>2+</sup> ) | PS+(DG+Ca <sup>2+</sup> ) |
| Tissue expression              | Universal                                     | Some tissues and cells | Many tissues and cells | Brain and spinal cord only   | Many tissues‡             | Brain only?‡              | Many tissues‡             |
| References (molecular cloning) | 5, 6, 9, 10, 13                               | 2, 7-11, 62, 63        | 2, 5-11                | 5-7, 9, 12, 13               | 4, 14                     | 4, 14, 15, 63             | 4, 14                     |

Additional cDNA clones probably related to the PKC family are partly sequenced (refs 56, 57). PS, phosphatidylserine; DG, diacylglycerol; AA, arachidonic acid.

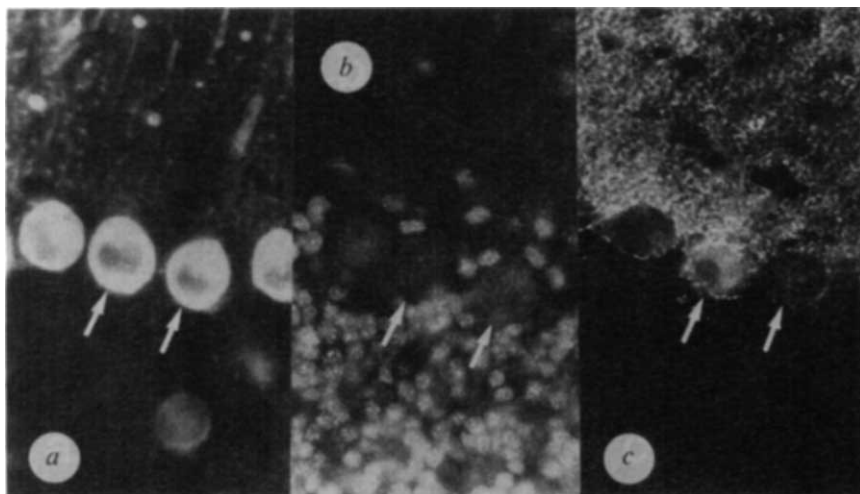
\* The complete sequence will be described elsewhere (Y. Ono *et al.*, manuscript in preparation).

† The activators for each subspecies are determined with calf thymus H1 histone as phosphate acceptor.

‡ The tissue expression is estimated only by Northern blot analysis (refs 4, 15).



**Fig. 3** Fluorescent micrographs of the rat cerebellar cortex stained with subspecies-specific antibodies against PKC. The upper portion illustrates the molecular layer, which contains the dendrites of the Purkinje cells, the main output neurons of the cerebellum. The lower portion shows the granule layer, which contains granule cells and other local interneurons. Arrows indicate Purkinje cell bodies. *a*, Stained with antibodies against the  $\gamma$  subspecies; *b*, Stained with antibodies against the  $\beta_I$ -subspecies. Granule cells are specifically stained but no immunofluorescence is detected in the molecular layer. *c*, Stained with antibodies against the  $\beta_{II}$ -subspecies. Under higher magnification, staining is observed to be present in terminal-like structures in the molecular layer and surrounding the Purkinje cell bodies. The detailed experimental conditions are described elsewhere<sup>31</sup>.



Immuno-electron microscopy shows that in some cells the  $\beta_{II}$ -subspecies is also associated with the Golgi membrane (C. Tanaka *et al.*, unpublished observation).

Many tissues, including liver, kidney, spleen, lung, heart and testis, contain the  $\beta$  subspecies (particularly  $\beta_{II}$ ) in variable ratios, but the  $\alpha$ -subspecies of PKC is the most widely distributed<sup>23,26,32-34</sup>. Some tissues such as heart, lung, adrenal cortex and platelets, appear to contain several undefined subspecies<sup>32,36</sup>. Most cell types contain more than one subspecies of PKC. For instance, we have recently observed that T lymphocytes and various clonal cell lines express the  $\alpha$ -,  $\beta_I$  and  $\beta_{II}$  subspecies in different ratios; their intracellular distribution may depend on the state of proliferation of the cells<sup>37</sup> (see also ref. 38).

### Individual characteristics

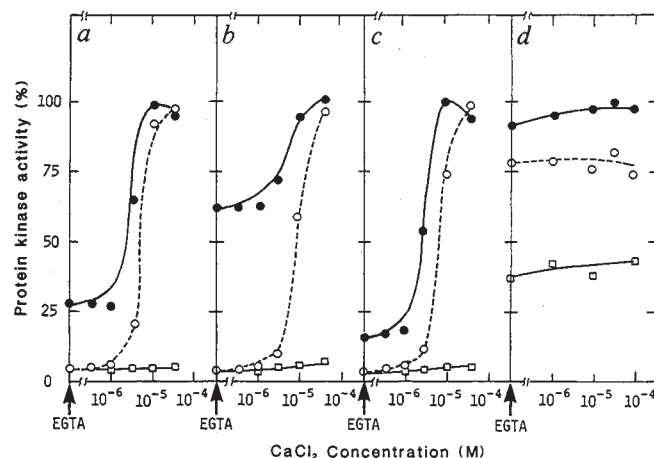
If the various PKC subspecies have different biological roles, one might expect some differences in their enzymatic properties. We already have some information on this from studies of chromatographically fractionated PKC activities<sup>39-44</sup>. Three activities (types I, II and III) fractionated on hydroxyapatite columns have been shown to correspond to  $\gamma$ ,  $\beta_I$  plus  $\beta_{II}$  and  $\alpha$  subspecies, respectively<sup>3,26</sup>. It remains to be seen how the more recently identified subspecies are fractionated on such columns.

It is already clear that the activities of  $\gamma$ ,  $\beta$  and  $\alpha$  are subtly different. Some of the kinetic variations are illustrated in Fig. 4. Thus, the  $\gamma$  (Fig. 4a) and  $\alpha$ -subspecies (Fig. 4c) are much less activated by diacylglycerol in the presence of phosphatidylserine than is the mixture of  $\beta_I$  and  $\beta_{II}$  (Fig. 4b), which shows substantial activity in the absence of  $\text{Ca}^{2+}$ . (The  $\beta_I$  and  $\beta_{II}$  subspecies show nearly identical kinetic properties, and cannot be separated from each other by this column chromatography<sup>2</sup>.) By contrast the  $\beta$  subspecies are poorly activated by free arachidonic acid, whereas the  $\gamma$  subspecies is activated by micromolar arachidonic acid (Z. Naor *et al.*, unpublished observation; see also refs 42, 45) and the  $\alpha$ -subspecies responds to high concentrations, but only at elevated concentrations of  $\text{Ca}^{2+}$  (refs 40, 42, 46). It is possible that different subspecies of PKC are activated by the series of phospholipid metabolites, such as diacylglycerol, arachidonic acid and lipoxin A<sup>47</sup>, that are produced in successive phases of the response of the cell to stimulation of a cell-surface receptor.

The structurally undefined enzymes from tissues such as heart, lung and platelets respond to phospholipid, diacylglycerol and  $\text{Ca}^{2+}$  in different ways<sup>32</sup>. For instance, one human platelet PKC is not sensitive to  $\text{Ca}^{2+}$  (Fig. 4d). This enzyme activity is reminiscent of the previously described rabbit reticulocyte PAKII protein kinase, which can phosphorylate the ribosomal S6 protein<sup>48</sup>,

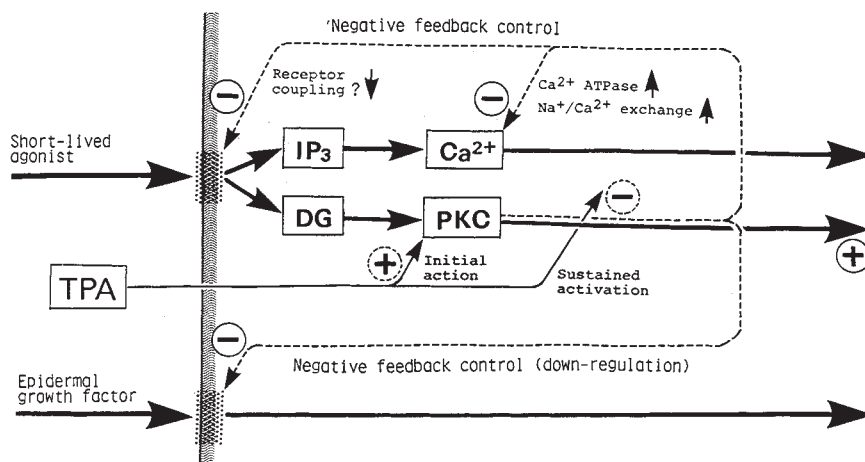
one of the targets of PKC. The exact relation of this enzyme to the members of the PKC family described above remains unclear.

Note that the *in vitro* dependency of PKC on  $\text{Ca}^{2+}$ , diacylglycerol and phospholipid varies markedly with the phosphate acceptor<sup>49</sup>, making it possible to compare the members of the PKC family under only limited, defined conditions. It is a general problem in protein phosphorylation research that in cell-free systems most protein kinases can phosphorylate many proteins (reviewed in ref. 50), only some of which are physiological substrates<sup>51</sup>. Thus, the possibilities that the enzymes respond differently to activators *in vivo* and that PKC subspecies are differentially distributed in various cell types and even within the same cell, greatly complicate the search for physiological substrates for the members of the PKC family. Presumably,



**Fig. 4** Kinetic variation in the activation of several subspecies of PKC. Activation is by phosphatidylserine and diacylglycerol in the presence of various concentrations of  $\text{CaCl}_2$ . Each subspecies was assayed with calf thymus  $\text{H}_1$  histone as a phosphate acceptor under the same conditions. Where indicated, PS ( $8 \mu\text{g ml}^{-1}$ ), DG ( $0.8 \mu\text{g ml}^{-1}$ ) and EGTA ( $0.5 \text{ mM}$ ) were added. PS, phosphatidylserine; DG, diolein; EGTA, ethylene glycol bis ( $\beta$ -aminoethyl-ether)- $N,N,N',N'$ -tetraacetic acid. Other detailed conditions are given elsewhere<sup>40,42</sup>.  $\bullet$ — $\bullet$ , in the presence of PS plus DG;  $\circ$ — $\circ$ , in the presence of PS alone;  $\square$ — $\square$ , in the absence of PS and DG. *a*, Rat  $\gamma$  subspecies (type I PKC). *b*, A mixture of rat  $\beta_I$  and  $\beta_{II}$  subspecies (type II PKC; these two subspecies are kinetically indistinguishable from each other). *c*, Rat  $\alpha$ -subspecies (type III PKC). *d*, A structurally undefined enzyme isolated from human platelets. More detailed properties of this enzyme will be described elsewhere (M. Tsukuda *et al.*, manuscript in preparation).

**Fig. 5** A model for the dual action of PKC and of TPA. This scheme emphasizes a negative feedback role of PKC in cellular regulation. Unlike short-lived diacylglycerol, TPA appears to show a positive action that initially activates PKC, then a negative action that initiates the degradation of the enzyme during sustained activation (see text), DG, diacylglycerol; IP<sub>3</sub>, inositol-1,4,5-trisphosphate.



some of the substrate preferences displayed by the different subspecies will result from the location of the substrate proteins in different intracellular compartments. Important questions therefore remain concerning the *in vivo* mechanism of activation of PKC and the true physiological target proteins for each subspecies.

### Dual action and implications

There have been several suggestions of possible functions for PKC, including involvement in secretion and exocytosis, modulation of ion conductance, regulation of receptor interaction with components of the signal transduction apparatus, smooth muscle contraction, gene expression and cell proliferation (reviewed in ref. 50). These diverse functions may well involve activation of different PKC subspecies. Although at present, there is little evidence of specific examples, studies on the biological roles of individual PKC subspecies will probably be prominent in the future.

It is now well recognized that the synergistic interaction between the PKC and Ca<sup>2+</sup> pathways, proposed earlier<sup>1</sup>, underlies a variety of cellular responses to external stimuli. The positive forward action of PKC seems to be important in the control of gene expression, such as induction of the interleukin-2 receptor and the activation of some proto-oncogenes (reviewed in ref. 50). Recent studies with fibroblasts over-expressing a PKC subspecies suggest that such actions are important in growth regulation<sup>52,53</sup>.

In addition to this positive forward action, a large body of evidence has now accumulated to indicate that PKC provides negative feedback control over various steps of the cell-signalling processes, operating in both short-term and long-term responses of the cell. In the short term, for example, PKC appears to be important in decreasing the inositol-1,4,5-trisphosphate (InsP<sub>3</sub>)-induced elevation of intracellular Ca<sup>2+</sup> levels, an effect that may occur at different points in the signalling pathway (Fig. 5). PKC may inhibit Ca<sup>2+</sup> mobilization by blocking the receptor-mediated hydrolysis of inositol phospholipids and hence production of InsP<sub>3</sub> and diacylglycerol, or by stimulating the hydrolysis of InsP<sub>3</sub> by activating an InsP<sub>3</sub> phosphatase<sup>54</sup>. Alternatively, PKC may stimulate the removal of intracellular Ca<sup>2+</sup> by activation of the Ca<sup>2+</sup>-transport ATPase and the Na<sup>+</sup>/Ca<sup>2+</sup> exchange protein (reviewed in ref. 50). The recent observations that 12-O-tetradecanoylphorbol-13-acetate (TPA) can inhibit the frequency of repetitive Ca<sup>2+</sup> transients recorded from single cells<sup>55</sup>, could be explained by an inhibitory action of PKC at the intracellular site of Ca<sup>2+</sup> release.

Although it sounds paradoxical, the negative feedback role of PKC may be extended to long-term responses, such as cell proliferation. The epidermal growth factor (EGF) receptor has repeatedly been shown to be phosphorylated by PKC, both *in*

*vivo* and *in vitro*. This results in a rapid decrease in high-affinity binding of EGF and inhibition of ligand-induced tyrosine phosphorylation, to give a functionally 'down-regulated' receptor (Fig. 5; reviewed in ref. 56). An analogous situation appears to exist with the T-cell receptor, present on T lymphocytes. Stimulation of these cells with TPA and Ca<sup>2+</sup> ionophore results in the phosphorylation and subsequent down-regulation of the receptor, preventing any further proliferative response to antigen<sup>57</sup>.

The dual action of PKC described above provides a versatile regulatory system that is finely tuned by the transient generation of second messengers such as diacylglycerol. It is clear, however, that agents that persistently activate the enzyme, such as TPA, can promote more complex and fundamental changes in its activity. As described above, TPA can stimulate the dual positive forward and negative feedback actions of PKC in a manner similar to diacylglycerol, and elicits many of the documented physiological effects of this enzyme. The secondary persistent action of TPA, which prolongs the association of PKC with the membrane, can, however, initiate the degradation of the PKC molecule and its sustained disappearance from the cell<sup>58,61</sup>. Our earlier studies<sup>18</sup> showed that calpain I, which is activated at the micromolar range of Ca<sup>2+</sup>, cleaves PKC in the presence of phosphatidylserine and diacylglycerol or TPA, implying that the activated form of PKC is a target of calpain action. Interestingly, several subspecies of PKC co-expressed in a single cell type have been found to disappear at different rates upon treatment with TPA<sup>61</sup>, which may reflect the substrate specificity of the calpain action (A. Kishimoto *et al.*, manuscript in preparation). Removal of PKC from the cell in this way would relieve the negative feedback control exerted by the enzyme, and, in the most extreme case, lead to uncontrolled cell proliferation in the presence of a mitogenic stimulus (Fig. 5). Although this hypothetical sequential action of tumour promoters—that is, a positive short-term activation of PKC, followed by a negative action, initiating the degradation of the enzyme over a longer time—has not yet been completely substantiated, it is an intriguing example of a way in which perturbation of the PKC pathway may have profound cellular effects.

Some of the experimental results reported from this laboratory were obtained in collaboration with the Biotechnology Laboratories of Takeda Chemical Industries and the Department of Pharmacology, Kobe University School of Medicine. Investigations in this laboratory were supported by the Research Fund of the Ministry of Education, Science and Culture, Japan; Muscle Dystrophy Association, US; Yamanouchi Foundation for Research on Metabolic Disorders; Merck Sharp and Dohme Research Laboratories; Ajinomoto Central Research Laboratories; Meiji Institute Health Sciences and New Lead Research Laboratories of Sankyo Company.



1. Nishizuka, Y. *Nature* **308**, 693-697 (1984).
2. Ono, Y. *et al. Science* **236**, 1116-1120 (1987).
3. Kikkawa, U. *et al. FEBS Lett.* **217**, 227-231 (1987).
4. Ono, Y. *et al. J. biol. Chem.* **263**, 6927-6932 (1988).
5. Parker, P. J. *et al. Science* **233**, 853-859 (1986).
6. Coussens, L. *et al. Science* **233**, 859-866 (1986).
7. Knopf, J. L. *et al. Cell* **46**, 491-502 (1986).
8. Ono, Y. *et al. FEBS Lett.* **206**, 347-352 (1986).
9. Kikkawa, U. *et al. FEBS Lett.* **223**, 212-216 (1987).
10. Ohno, S. *et al. Nature* **325**, 161-166 (1987).
11. Kubo, K., Ohno, S. & Suzuki, K. *FEBS Lett.* **223**, 138-142 (1987).
12. Ohno, S. *et al. Biochemistry* **27**, 2083-2087 (1988).
13. Ono, Y. *et al. Nucleic Acid Res.* **16**, 5199-5200 (1988).
14. Ono, Y. *et al. FEBS Lett.* **226**, 125-128 (1987).
15. Ohno, S., Akita, Y., Konno, Y., Imajoh, S. & Suzuki, K. *Cell* **53**, 731-741 (1988).
16. Berg, J. *Science* **232**, 485-487 (1986).
17. Hunter, T. & Cooper, J. A. *Rev. Biochem.* **54**, 897-930 (1985).
18. Kishimoto, A., Kajikawa, N., Shiota, M. & Nishizuka, Y. *J. biol. Chem.* **258**, 1156-1164 (1983).
19. Nagle, D. S. & Blumberg, P. M. *Cancer Lett.* **18**, 35-50 (1983).
20. Worley, P. F., Baraban, J. M. & Snyder, S. H. *J. Neurosci.* **6**, 199-207 (1986).
21. Brandt, S. J., Nijdel, J. E., Bell, R. M. & Young III, W. S. *Cell* **49**, 57-63 (1987).
22. Wood, J. G., Girard, P. R., Mazzei, G. J. & Kuo, J. F. *J. Neurosci.* **6**, 2571-2577 (1986).
23. Shearman, M. S., Naor, Z., Kikkawa, U. & Nishizuka, Y. *Biochem. biophys. Res. Commun.* **147**, 911-919 (1987).
24. Mochly-Rosen, D., Basbaum, A. J. & Koshland, D. E. Jr. *Proc. natl. Acad. Sci. U.S.A.* **84**, 4660-4664 (1987).
25. Huang, F. L., Yoshida, Y., Nakabayashi, H. & Huang, K.-P. *J. biol. Chem.* **262**, 15714-15720 (1987).
26. Huang, F. L. *et al. Biochem. biophys. Res. Commun.* **149**, 946-952 (1987).
27. Kitano, T. *et al. J. Neurosci.* **7**, 1520-1525 (1987).
28. Saito, N., Kikkawa, U., Nishizuka, Y. & Tanaka, C. *J. Neurosci.* **8**, 369-382 (1988).
29. Kose, A. *et al. J. Neurosci.* (in the press).
30. Hashimoto, T. *et al. J. Neurosci.* **8**, 1678-1683 (1988).
31. Ase, K. *et al. J. Neurosci.* (in the press).
32. Kosaka, Y. *et al. Biochem. biophys. Res. Commun.* **151**, 973-981 (1988).
33. McCaffrey, P. G. *et al. Biochem. biophys. Res. Commun.* **146**, 140-146 (1987).
34. Shearman, M. S., Kosaka, Y., Ase, K., Kikkawa, U. & Nishizuka, Y. *Biochem. Soc. Trans.* **16**, 307-308 (1988).
35. Hidaka, H. *et al. J. biol. Chem.* **263**, 4523-4526 (1988).
36. Pelosin, J., Vilgrain, I. & Chambaz, E. M. *Biochem. biophys. Res. Commun.* **147**, 382-391 (1987).
37. Shearman, M. S. *et al. FEBS Lett.* **234**, 387-391 (1988).
38. Makowski, M., Ballester, R., Cayre, Y. & Rosen, O. M. *J. biol. Chem.* **263**, 3402-3410 (1988).
39. Huang, K.-P., Nakabayashi, H. & Huang, F. L. *Proc. natl. Acad. Sci. U.S.A.* **83**, 8535-8539 (1986).
40. Sekiguchi, K., Tsukuda, M., Ogita, K., Kikkawa, U. & Nishizuka, Y. *Biochem. biophys. Res. Commun.* **145**, 797-802 (1987).
41. Jaken, S. & Kiley, S. C. *Proc. natl. Acad. Sci. U.S.A.* **84**, 4418-4422 (1987).
42. Sekiguchi, K., Tsukuda, M., Ase, K., Kikkawa, U. & Nishizuka, Y. *J. Biochem. (Tokyo)* **103**, 759-765 (1988).
43. Asaoka, Y. *et al. FEBS Lett.* **231**, 221-224 (1988).
44. Ido, M., Sekiguchi, S., Kikkawa, U. & Nishizuka, Y. *FEBS Lett.* **219**, 215-218 (1987).
45. Kikkawa, U. *et al. Cold Spring Harbor Symposia on Quant. Biol.* **53** (in the press).
46. Wooten, M. E. & Wrenn, R. W. *Biochem. biophys. Res. Commun.* **153**, 67-73 (1988).
47. Hansson, A. *et al. Biochem. biophys. Res. Commun.* **134**, 1215-1222 (1986).
48. Gonzatti-Haces, M. I. & Traugh, J. A. *J. biol. Chem.* **261**, 15266-15272 (1986).
49. Bazzi, M. D. & Nelsestuen, G. L. *Biochemistry* **26**, 1974-1982 (1987).
50. Nishizuka, Y. *Science* **233**, 305-312 (1986).
51. Krebs, E. G. & Beavo, J. A. *Rev. Biochem.* **48**, 923-959 (1979).
52. Housey, G. M. *et al. Cell* **52**, 343-354 (1988).
53. Persons, D. A., Wilkinson, W. O., Bell, R. M. & Finn, O. J. *Cell* **52**, 447-458 (1988).
54. Connolly, T. M., Lawing, J. F. & Majerus, P. W. *Cell* **46**, 951-958 (1986).
55. Woods, N. M., Cuthbertson, K. S. R. & Cobbold, R. H. *Biochem. J.* **246**, 619-623 (1987).
56. Schllessinger, J. *J. Cell Biol.* **103**, 2067-2072 (1986).
57. Cantrell, D. A., Davies, A. A. & Crumpton, M. J. *Proc. natl. Acad. Sci. U.S.A.* **52**, 8158-8162 (1985).
58. Young, S., Parker, P. J., Ullrich, A. & Stabel, S. *Biochem. J.* **244**, 775-779 (1987).
59. Fournier, A. & Murray, A. W. *Nature* **330**, 767-764 (1987).
60. Young, S., Rothbard, J. & Parker, P. J. *Eur. J. Biochem.* **173**, 247-252 (1988).
61. Ase, K., Berry, N., Kikkawa, U., Kishimoto, A. & Nishizuka, Y. *FEBS Lett.* (in the press).
62. Coussens, L., Rhee, L., Parker, P. J. & Ullrich, A. *DNA* **6**, 389-394 (1987).
63. Housey, G. M., O'Brain, C. A., Johnson, M. D., Kirshemeier, P. & Weinstein, I. B. *Proc. natl. Acad. Sci. U.S.A.* **84**, 1065-1069 (1987).
64. Makowski, M., Birnbaum, M. J., Ballester, R. & Rosen, O. M. *J. biol. Chem.* **261**, 13389-13392 (1986).

## ARTICLES

## Global fire at the Cretaceous-Tertiary boundary

Wendy S. Wolbach\*, Iain Gilmour\*, Edward Anders<sup>†‡</sup>, Charles J. Orth<sup>‡</sup>  
& Robert R. Brooks<sup>§</sup>

\* Enrico Fermi Institute and Department of Chemistry, University of Chicago, Chicago, Illinois 60637-1433, USA

† Physikalisches Institut der Universität, 3012 Bern, Switzerland

‡ Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

§ Departments of Chemistry and Biochemistry, Massey University, Palmerston North, New Zealand

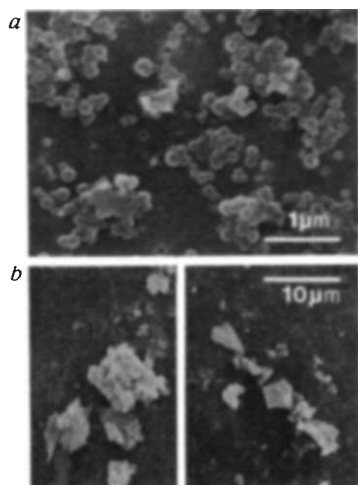
*Cretaceous-Tertiary (K-T) boundary clays from five sites in Europe and New Zealand are 10<sup>2</sup>-10<sup>4</sup>-fold enriched in elemental C (mainly soot), which is isotopically uniform and apparently comes from a single global fire. The soot layer coincides with the Ir layer, suggesting that the fire was triggered by meteorite impact and began before the ejecta had settled.*

IN an earlier paper<sup>1</sup> we reported the presence of soot at three widely separated K-T boundary sites, and suggested that it came from major fires triggered by the impact of a giant meteorite. Though the impact theory<sup>2,3</sup> has been disputed<sup>4,5</sup>, it alone seems able to account for the full range of evidence<sup>6</sup>, including the data in ref. 1 and this paper. In an attempt to resolve the sequence of events at the K-T boundary, we have made a detailed study of the Woodside Creek, New Zealand, site<sup>7</sup>, whose uniquely sharp Ir profile and high Ir abundance<sup>8</sup> imply minimal disturbance. Apart from C, we have measured Ir and 32 other elements by RNAA and INAA as well as carbon and oxygen isotopes (W.S.W. *et al.*, in preparation). Our samples were portions of the powders used by Brooks *et al.*<sup>8</sup>. We also measured several samples from Chancton Rocks, New Zealand<sup>9,10</sup>, Stevns Klint, Denmark, and Nye Kløv, Denmark. The Danish samples were kindly provided by H. J. Hansen of the University of Copenhagen.

The K-T boundary at Woodside Creek has been described by Strong<sup>7</sup>. The boundary is marked by a layer of iron-rich clay ~6-8 mm thick, separating thickly bedded limestones of late

Maastrichtian age from the more thinly bedded siliceous limestones of lower Teurian age. The rocks have experienced little structural modification<sup>7</sup> and probably represent an upper bathyal environment with water depths of ~600-800 m. The boundary clay, in particular, seems undisturbed, as both it and the first 2 cm of overlying limestone are laminated on a <1-mm scale.

A major improvement over our previous work was a new procedure for resolving the two principal carbon components, kerogen and elemental carbon, and for estimating the proportions of soot (Fig. 1a) and coarse carbon (Fig. 1b) in the latter. Earlier authors<sup>11</sup> had concluded that similar coarse carbon in marine sediments is charcoal derived from fires on land. As before<sup>1</sup>, we isolated a carbonaceous residue by dissolving the rock in HF-HCl, but then removed kerogen by etching with Cr<sub>2</sub>O<sub>7</sub><sup>2-</sup> under appropriate conditions (ref. 12 and W.S.W. and E.A., in preparation), rather than by NaOH and organic solvents<sup>1</sup>. The unetched and etched samples were then combusted to CO<sub>2</sub> for mass-spectrometric analysis, yielding isotopic data and weights for elemental C and kerogen (the latter by



**Fig. 1** Morphology of K-T carbon. *a*, Soot of characteristic 'grape bunch' morphology dominates in boundary clay (0.3–0.4 cm above boundary). *b*, Irregularly shaped, platey or pitted coarse carbon, presumably charcoal, dominates at higher levels (17.4–19.9 and 3.6–5.8 cm). Unlike soot, which forms only in flames<sup>48</sup>, charcoal forms at lower temperatures, by charring and devolatilization of fuel. All samples were etched with dichromate to remove kerogen.

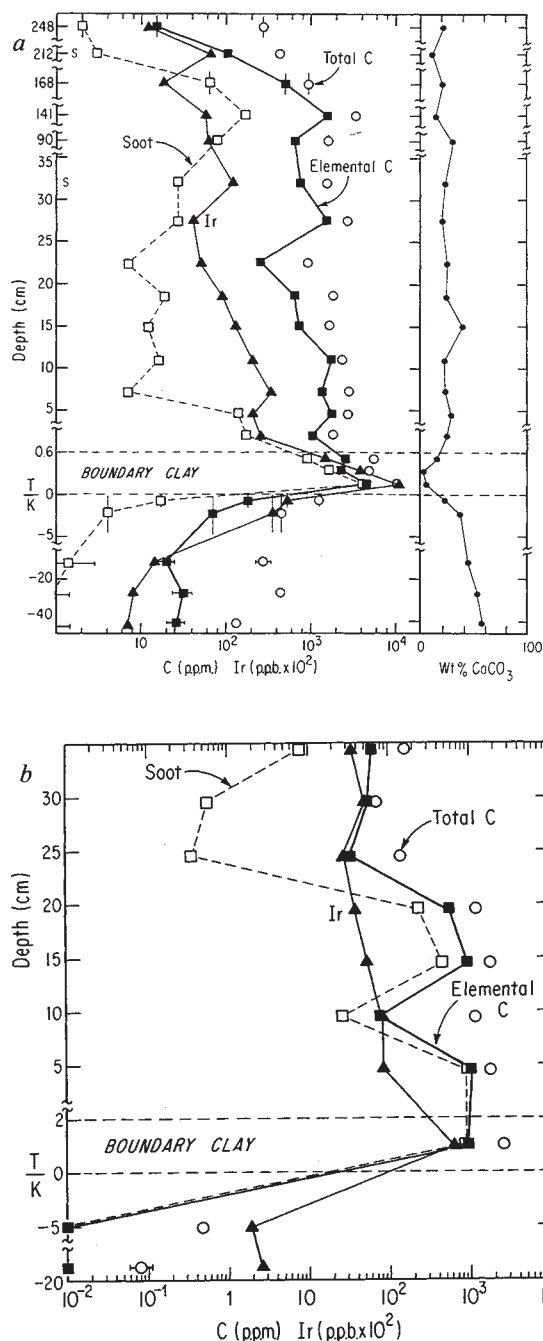
difference) as well as the content of insoluble minerals. The elemental C was examined by scanning electron microscope (SEM) and the mass fraction of soot was found by planimetric analysis<sup>13</sup>.

### Carbon and iridium profiles at boundary

Figure 2*a* shows the data for the ~3-m section straddling the boundary, with the -5 to +35 cm and especially the 0 to +0.6 cm intervals expanded. No corrections for carbonate content were made. Iridium and elemental C concentrations both rise sharply at the boundary, by factors of 1,500 and 210 compared with the Cretaceous average of 0.076 parts per  $10^9$  (p.p.b.) and 22 p.p.m. Both reach their maximum in the basal 0.3 cm of the boundary clay and decline afterwards, but the decline for Ir is greater and more regular. Soot, with a Cretaceous average of  $1 \pm 0.5$  p.p.m., shows an even steeper rise ( $\times 3,600$ ) in the basal layer and becomes a major component (up to 71%) of the elemental C in the boundary clay, as shown by the convergence of the two curves. It then declines sharply, becoming a minor (<1 to 16%) component of the elemental carbon. All three curves revert to Cretaceous levels only at ~250 cm.

The large, abrupt rise in C and Ir at Woodside Creek is not unique. The Chancet Rocks site, 15 km away, shows a similar increase (Fig. 2*b*). Although this site is tectonically disturbed<sup>10</sup> and lacks boundary clay in places, it too shows a steep rise in elemental C and Ir at the boundary, by factors of  $3 \times 10^5$  and 290 respectively.

The rise in C at Chancet Rocks may have been exaggerated by local geological factors. Still, it is curious that our earlier data for the Danish K-T boundary—the only place for which we had comparison samples for the Cretaceous and Tertiary—showed only a fourfold increase in carbon abundance<sup>1</sup>. Two likely reasons are incomplete removal of kerogen and lack of a correction for insoluble minerals (mainly rutile), which would cause our carbon values to be high. We have therefore reanalysed the Nye Kløv marls, using our new procedure for removal of kerogen. The new elemental carbon values are substantially lower for both samples (Cretaceous 77 against 950 p.p.m., Tertiary 430 against 1,040 p.p.m.), corresponding to a 47-fold rise at the boundary (3,600 p.p.m.). The soot content rises even more sharply across the boundary: 111,930 and 97 p.p.m. Thus even the Danish samples show a ~180-fold increase in soot con-



**Fig. 2** Carbon and Ir profiles at the K-T boundary. *a*, At Woodside Creek the concentrations of Ir, elemental C and especially soot rise steeply at the boundary, by factors of 1,500, 210 and 3,600. Soot appears in the first 0.3 cm interval of the boundary clay, showing that fires started before the fallout had settled. After an initial drop above the boundary, all three decline gradually, returning to Cretaceous levels only at ~250 cm. This delayed fallout appears to be secondary, redeposited by lateral transport in topographic lows. Soot is the major form of elemental C in the basal layer of the boundary clay, but becomes much less abundant at higher levels. Vertical bars indicate depth range of discontinuous samples. S, shale partings, which form upon changes in deposition conditions. Soot values outside the boundary clay are uncertain by factors of 2–3. Carbonate was determined by decomposition with  $H_3PO_4$ . *b*, At Chancet Rocks, both Ir and elemental C show their characteristic rise at the K-T boundary, by factors of 290 and  $3 \times 10^5$ . The 2-cm boundary clay sample plotted here actually comes from another location 50 m away.



centration at the boundary. Such an increase cannot<sup>14</sup> possibly be due to condensation of the section by dissolution of carbonate, as argued by Officer and Ekdale<sup>15</sup>, because, as they themselves admit, this mechanism can raise concentrations at most by the ratio of clay contents, that is, a (constant) factor of 7 for the Danish section<sup>15</sup>. But the factors actually required are larger and highly variable: 47 for elemental C, 180 for soot and 2,000 for Ir (this work and refs 14, 16).

### Prompt or delayed fires?

Several authors (for example Argyle<sup>17</sup>) have pointed out that an impact is unlikely to start a global wildfire, as living trees do not burn well. They suggest instead that forests were killed and freeze-dried by the initial darkness and cold following the impact, and were then ignited by lightning after the sky had cleared and thunderstorms had resumed.

According to this model, soot should first appear on top of the Ir-bearing ejecta, as fires would start only after all ejecta had settled (in <1 yr; ref. 18). But Fig. 2a shows that soot accompanied Ir even in the first fallout fraction; both rise more than 200-fold in the bottom 0.3 cm of the boundary clay and then decline. Apparently the fire started well before all the ejecta had settled.

### Deposition time

First, let us check how sharp a time marker the Ir spike represents, using a plot that also includes three other elements enriched in boundary clay, all on a carbonate-free basis (Fig. 3). The Ir concentration rises in the -4.5 to 0 and -1.5 to 0 cm samples, but actually Al, Fe and Sb, which are non-meteoritic indicators of boundary clay, also rise in these two samples, and the carbonate content drops (Fig. 3). Thus the rise in Ir concentration in this interval apparently represents an admixture of ~5% boundary clay, not a premonitory Ir signal. (The scale length of mixing must be less than a few millimetres as the boundary clay itself shows lithological variations and large isotopic and chemical differences over millimetre distances (Figs 2a, 3, 4 and unpublished data).) Thus it seems that the Ir peak at 0 cm is indeed a sharp time marker, representing the sudden injection of  $\sim 10^{-7}$  g cm<sup>-2</sup> of Ir or  $\sim 0.2$  g cm<sup>-2</sup> of meteoritic material (if the Ir is meteoritic, for which there is ample evidence<sup>6,14,19</sup>).

Next we consider the settling time of the boundary clay. This clay, averaging 2–5 g cm<sup>-2</sup> worldwide<sup>2,3,6</sup>, apparently represents impact ejecta that subsequently weathered to smectite clay<sup>16</sup>, along with some local material. The original particle size of the ejecta may have been mainly >1  $\mu$ m, judging from some surviving material in the boundary clay: spherules up to 1000  $\mu$ m (refs 20, 21) and shocked quartz grains in the 1–100- $\mu$ m range<sup>22</sup>. These ejecta cannot have stayed aloft for more than 6 months<sup>18</sup>, and by Stokes's Law should have settled through 600–800 m of water in a similarly short time. The situation is analogous to volcanic ash, for which settling times of weeks or months are generally accepted<sup>23,24</sup>.

### Primary and secondary fallout

In the Tertiary concentrations of Ir and all three forms of C remain elevated above Cretaceous levels to at least +34 cm (Fig. 2), reverting to normal only above +213 cm at Woodside Creek. A similar trend for Ir alone has been seen at other sites<sup>2,19,25,26</sup>, and is usually interpreted in terms of two kinds of fallout: primary, represented by the boundary clay<sup>25</sup> or even only its basal layer<sup>26</sup>, and secondary, represented by the remaining material of elevated Ir content. Both kinds of fallout apparently come from a single impact, as they contain meteoritic and non-meteoritic elements (Ir, Ni and Cr; Sb and As) in characteristic proportions<sup>27</sup>. The amount of Ir and C in this secondary fallout is much larger than that in the primary fallout, as shown by the totals 0–0.6, 0–34 and 0–250 cm: Ir (91, 290 and 560 ng cm<sup>-2</sup>), total elemental C (4.8, 89 and 395 mg cm<sup>-2</sup>), and

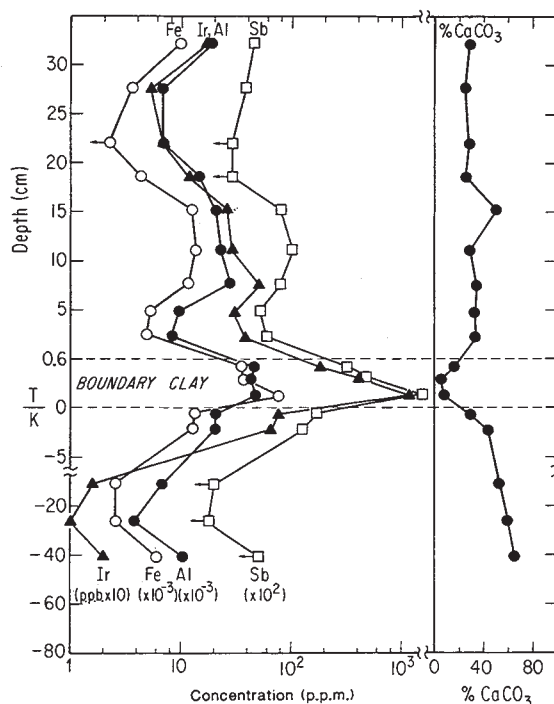


Fig. 3 Elements enriched in boundary clay. In the interval -4.5 to 0 cm, meteoritic Ir rises, but so do non-meteoritic Fe, Al and Sb, which are all good indicators of boundary clay, whereas the carbonate content drops. Apparently the rise in Ir content represents admixture of ~5% boundary clay to the 0 to -4.5 cm and 0 to -1.5 cm samples, not a premonitory Ir signal. Carbonate values were estimated from Ca contents.

soot (3.6, 19 and 20 mg cm<sup>-2</sup>). The time interval represented by 250 cm is of the order  $10^5$  yr, for an assumed sedimentation rate of 1–2 cm kyr<sup>-1</sup>.

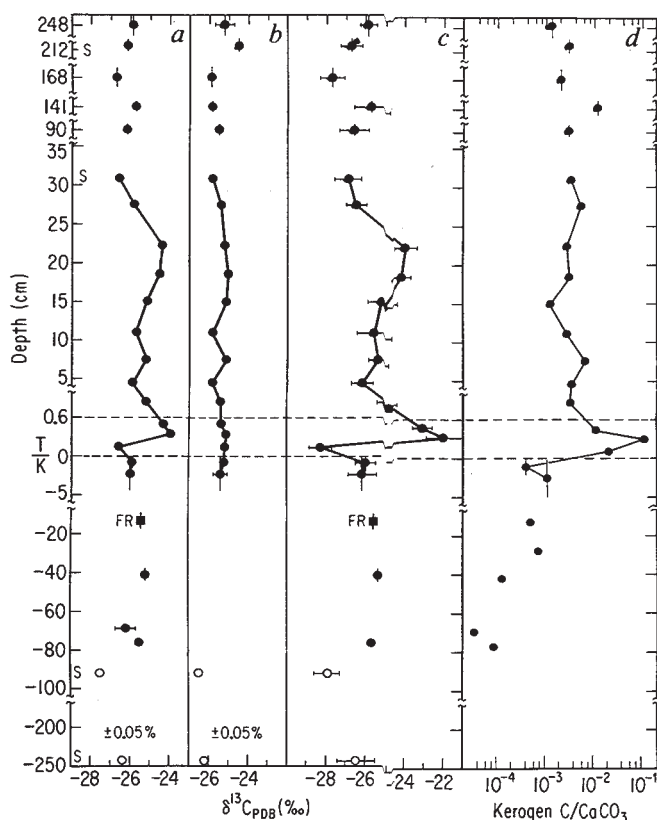
The amount of secondary Ir is too large to be runoff from land, although New Zealand was conterminous with Australia and Antarctica. More probably, secondary Ir is fallout eroded from elevated sea-floor sites that was redeposited in topographic lows. Such topographic lows have the best chance of retaining their primary fallout, but at the price of becoming sinks for secondary fallout. Only the primary fallout counts for our purposes.

Total elemental carbon shows an even greater tendency to concentrate in the secondary fallout (Fig. 2a):  $\log ([C_{elem}]/[Ir])$  rises steadily in the three layers of the boundary clay (4.6, 4.8, 5.2), levels off at  $5.7 \pm 0.1$  in the next seven intervals (to +25 cm), where the curves run roughly parallel, and then rises to  $6.2 \pm 0.3$  before declining again above +213 cm. Soot, on the other hand, correlates with Ir only from 0 to 3.6 cm ( $\log ([C_{soot}]/[Ir]) = 4.7 \pm 0.1$ ), dropping sharply and remaining fairly constant to 34 cm, while Ir declines. A tentative interpretation, in the spirit of a working hypothesis, is that coarse C and especially soot coagulated with the Ir-bearing ejecta and fell out together in the primary stage. During the lengthy secondary stage, soot was depleted by sorting, whereas coarse C stayed with the ejecta.

### Carbon isotopes

Figure 4 shows carbon isotope data for total and elemental carbon as well as kerogen, which was determined by difference and therefore has larger errors.

The elemental carbon from -4.5 to +248 cm is remarkably constant ( $-25.36 \pm 0.36\%$ ) and is about 1% heavier than its counterpart in the Cretaceous ( $-26.48$  and  $-26.16\%$ ). This constancy supports our earlier conclusion that all this carbon is global fallout from the K-T event. Moreover, there seems to be little isotopic difference between soot and coarse carbon, as



**Fig. 4** Isotopic data for reduced carbon, Woodside Creek. Total C (a) shows larger variations, especially in the boundary clay, which must be inherent in the kerogen. Elemental C (b) from -4.5 to +248 cm has nearly constant  $\delta^{13}\text{C}$ , supporting the interpretation of this carbon as fallout from a single source. For kerogen (c), the shift to -29‰ at 0–0.3 cm may be due to respiration of plankton in darkness (which preferentially eliminates  $^{13}\text{C}$ ), and the subsequent shift to -22‰ may reflect a rapid sweeping out and burial of plankton and other particulate organic carbon, without the usual bacterial degradation that preferentially consumes isotopically heavy proteins and carbohydrates. Indeed, the ratio kerogen C/CaCO<sub>3</sub> (d) rises ~10-fold in the boundary clay, suggesting a lapse in the preferential destruction of organic C. FR, a 164 g sample of Cretaceous limestone from the nearby Flaxbourne River site, which helps define the Cretaceous baseline. Open symbols, shale partings. Vertical bars indicate depth range of discontinuous samples.

samples differing widely in soot content have virtually identical isotopic compositions (for example, the intervals 0–0.3 and 5.8–9.3 cm, with 91% and 0.5% soot, have  $\delta^{13}\text{C} = -25.17\%$  and  $-25.12\%$ ).

The baseline for Cretaceous kerogen is not well defined (Fig. 4), as the total carbon content is low and the elemental carbon content even lower, often too little for isotopic measurement. We shall therefore ignore the variations in the Tertiary, which are comparable with those in the Cretaceous, and confine ourselves to the large variations in the boundary clay: a 3‰ swing to lighter values followed by a 7‰ rebound to heavier values. These are paralleled by an increase in aromatic character of the kerogen.

The 3‰ swing to -29‰ (Fig. 4) at 0–0.3 cm resembles that at +27 and thus might have the same commonplace cause, if it did not occur over a far shorter depth interval and especially time interval (<0.2 yr). A similar shift (-1.5‰) has been seen in dinocysts at Stevns Klint<sup>28</sup>. It may represent the first effects of the impact on the marine ecosystem, that is, darkness and acid rain<sup>29</sup>. Plankton may have lingered on by respiration, perhaps becoming isotopically lighter owing to a poison-induced change in respiratory substrate<sup>30</sup>. Alternatively, if death was

instantaneous, the shift may have been caused by abiotic reactions with acid rain<sup>29</sup> ( $\text{HNO}_2 + \text{HNO}_3$ ).

The 7‰ shift to heavier values in the next 0.1 cm, confirmed by another heavy value in the following 0.2 cm, is remarkable in terms of its magnitude and speed. A possible explanation is rapid burial of dead plankton with greatly reduced alteration by oxidation and bacterial degradation. Normally proteins, nucleic acids and sugars—all of which are enriched<sup>31</sup> in  $^{13}\text{C}$ —are preferentially destroyed during descent, leaving mainly the less reactive, light components (lipids, cellulose). But the massive amounts of K-T ejecta (2–5 g cm<sup>-2</sup>, with ~10<sup>3</sup> cm<sup>2</sup> cross-sectional area per cm<sup>2</sup> of ocean surface) settling rapidly through the water column would sweep along fresh plankton and other particulate organic carbon (POC), thus protecting it from bacterial attack first by rapid descent and then by burial.

Direct evidence for such mass death and rapid burial has been obtained at Stevns Klint, where the number of dinocysts increases ~10<sup>4</sup>-fold from -3 to +0.5 cm, peaking just above the rusty basal layer of the boundary clay<sup>28</sup>. Further evidence is available at Woodside Creek. (1) The kerogen concentration in the boundary clay, 2,000 p.p.m., is ~15-fold higher than between -80 and -10 cm. (2) The ratio kerogen/CaCO<sub>3</sub>, two biogenic materials, peaks sharply in the boundary clay, exceeding late Cretaceous and early Tertiary values by ~100-fold and ~30-fold (Fig. 4d). (3) The total amount of kerogen in the boundary clay, 0.0057 g C cm<sup>-2</sup>, is comparable with the global average of POC (living and dead) in the present oceans, 0.0014 g C cm<sup>-2</sup> (refs 32, 33). Under normal deposition conditions, only ~1% of the original POC flux survives oxidation and biodegradation and ends up as kerogen, but in the boundary clay essentially all did. (4) In the basal layer of the boundary clay, N is 20-fold enriched over Cretaceous or Tertiary values, and the N/C ratio is 0.20, compared with 0.03–0.05 in most of the section<sup>34</sup>. Such an enhanced abundance of N is consistent with enhanced preservation of nitrogen compounds, perhaps augmented by nitrosylation by HNO<sub>2</sub> from acid rain.

## Biomass or fossil carbon

Few tests are available for choosing between these two potential sources of K-T carbon<sup>1</sup>. A high potassium content is diagnostic of wood soot<sup>35</sup>, but we found that the potassium in wood soot is readily leached by the acids used to dissolve the sediment, dropping from 24,000 to 0.3 p.p.m. after 12 h at 65 °C. Morphology is equally uninformative, as the coarse K-T carbon (Fig. 1b) contains no spheres or highly elongated particles that are diagnostic of oil or wood<sup>36</sup>. However, the K-T particles are much smaller than those studied by Griffin and Goldberg<sup>36</sup> (~10 μm compared with >38 μm but typically ~100 μm), and as they come from a very much larger fire it is not clear that the absence of characteristic shapes is significant.

A strong argument for biomass comes from polynuclear aromatic hydrocarbons in the sediment. K-T boundary clays from Woodside Creek (I.G. and F. Guenther, in preparation) and the North Atlantic DSDP site 605 (ref. 37) contain the tricyclic hydrocarbon retene (methyl-7-isopropyl phenanthrene), which may be diagnostic of coniferous wood combustion<sup>38</sup>. (It probably forms by mild thermal degradation of the diterpenoid resin acid, abietic acid<sup>38</sup>.) Also, the  $\delta^{13}\text{C}$  value of K-T carbon,  $-25.4 \pm 0.3\%$ , agrees best with that for natural charcoal from forest fires. In the compilation by Deines<sup>31</sup>, 43% of the charcoal samples fall between -25 and -26‰, compared with only 16, 18 and 4% of C<sub>3</sub> plants, coal or oil. As the  $\delta^{13}\text{C}$  distributions for charcoal and C<sub>3</sub> plants are 8–14‰ wide, the remarkable uniformity of the K-T carbon ( $\pm 0.3\%$ ; Table 1) implies global mixing of the combustion products before fallout.

The global amount of K-T carbon,  $(7 \pm 4) \times 10^{16}$  g, is very large: about 10% of the present (above-ground) biomass carbon<sup>39</sup> or 3% of the maximum Cretaceous carbon (calculated on the extreme assumption that all land was covered with biomass at the highest known density, 1.5 g C cm<sup>-2</sup>, as for



**Table 1** Elemental carbon in K-T boundary clay at five sites

| Site                  | Carbon abundance<br>(g cm <sup>-2</sup> ) | $\delta^{13}\text{C}_{\text{PDB}}^*$<br>(‰) |
|-----------------------|-------------------------------------------|---------------------------------------------|
| Woodside Creek, NZ    | 0.0048                                    | -25.23                                      |
| Chancet Rocks, NZ     | 0.025                                     | -25.42                                      |
| Stevns Klint, Denmark | 0.011                                     | -25.81                                      |
| Caravaca, Spain       | 0.010                                     | -25.00                                      |
| Gubbio, Italy         | 0.013                                     | -25.48                                      |
| Mean                  | 0.013 $\pm$ 0.007                         | -25.39 $\pm$ 0.30                           |

\* All errors  $\pm 0.05\%$ .

tropical rain forests<sup>39</sup>). Thus 3% of the living carbon would have to be converted to soot if all the Earth's forests burned down and even more if only some of them did. In present-day forest fires<sup>40,41</sup>, soot yields range from 0.1 to 2%, and although they may be higher in large fires where oxygen access is limited<sup>1</sup>, this factor is offset by the difficulty of burning down all of Earth's forests. Natural firebreaks will protect some areas, and natural moisture will protect others. A higher partial pressure of oxygen, as little as 24 vol.% compared with the 30% inferred for the Late Cretaceous<sup>42</sup>, would enable even moist wood to ignite<sup>43</sup>, but, with daily thunderstorms (the present frequency is  $\sim 10^4$  per day), wildfires would be raging all the time, burning down forests as fast as they grew.

Perhaps the most likely possibility is that trees were killed and dried by the impact, both by the prompt heating of the atmosphere and strong winds capable of flattening forests out to a distance of 500–1000 km<sup>44</sup>, and subsequent heating by the ejecta plume and hot fallout. The kinetic energy of the projectile,  $10^{31}$  ergs, corresponds to  $5 \times 10^4$  cal cm<sup>-2</sup>, and even a small fraction of this would suffice to start major wildfires. If the impact was in North America<sup>45</sup>, then the fires might spread over Eurasia and perhaps Africa, before being stopped by the ocean.

The fires may also have ignited some fossil carbon, thus augmenting the biomass source. Most of it is buried and thus would not ignite unless exposed by the impact, but in view of the high mean abundance of fossil C, 2,300 g cm<sup>-2</sup>, even the carbon exposed by a 200–300-km crater in average terrain would make a significant contribution. Apart from oil and coal<sup>1</sup>, the Cretaceous black shales<sup>46</sup> have been proposed as possible sources.

## Events at the K-T boundary

A global fire producing  $7 \times 10^{16}$  g of soot would aggravate most of the environmental stresses of an impact<sup>2</sup>. As soot absorbs

sunlight more effectively and settles more slowly than does rock dust, the darkness and cold would last longer. (The optical depth of 0.013 g cm<sup>-2</sup> of soot is about 1,800, before coagulation.) Poisons such as NO and NO<sub>2</sub> (ref. 47) would be accompanied by CO ( $\sim 100$  p.p.m.<sup>41</sup>) and organic pyrotoxins<sup>41</sup>. The latter, in particular, may help explain the selectivity of extinction patterns, which some authors<sup>3</sup> have claimed to be inconsistent with the non-specific stresses of an impact, as different species often have very different tolerances for chemical toxins. Lastly, the CO<sub>2</sub> accompanying the soot ( $\sim 900$  p.p.m.<sup>41</sup>) would increase the greenhouse effect due to water vapour ( $\sim 8^\circ\text{C}$ ; ref. 44) by another  $\sim 9^\circ\text{C}$ .

Evidently K-T boundary clay contains a record of the impact and its aftermath, which is still preserved at relatively undisturbed sites such as Woodside Creek. Our data suggest the following sequence of events in the three layers studied.

**0–0.3 cm.** Ir concentration increases abruptly by a factor of 1,500 over the Cretaceous, due to meteorite impact and rapid settling of ejecta. Elemental carbon (210 $\times$ ) and soot (3,600 $\times$ ) increase concurrently, because of major fires triggered by the impact. Kerogen increases >15-fold, presumably due to a partial sweeping out of live plankton near surface and dead organic matter at depth. Carbonate content drops from 44 to 5%. Carbonate  $\delta^{13}\text{C}$  becomes 1.3‰ lighter and  $\delta^{18}\text{O}$  becomes 1.8‰ lighter. Kerogen  $\delta^{13}\text{C}$  becomes 3‰ lighter, possibly due to respiration in darkness or reactions with acid rain.

**0.3–0.4 cm.** Abundances decline relative to the basal layer: Ir (0.35), elemental carbon (0.50), soot (0.40), kerogen (0.45), carbonate (0.67). Carbonate  $\delta^{13}\text{C}$  becomes heavier by 0.9‰;  $\delta^{18}\text{O}$  by 1.3‰, and kerogen  $\delta^{13}\text{C}$  by 7‰ (rapid burial without biodegradation).

**0.4–0.6 cm.** Abundances continue to drop relative to the basal layer: Ir (0.14), elemental C (0.54), soot (0.22) and kerogen (0.31), whereas carbonate increases ( $\times 2.9$ ). Kerogen  $\delta^{13}\text{C}$  remains 5‰ heavier (continued rapid burial without biodegradation), and  $\delta^{18}\text{O}$  becomes  $\sim 1.0\%$  lighter.

The chemical and isotopic differences among the three layers show that the scale length for mixing was  $<0.1$  cm in the boundary clay at this site, permitting a detailed stratigraphic record to survive. The data suggest a series of environmental changes rather like those expected from a meteorite impact<sup>2</sup> followed by major fires<sup>1</sup>.

We thank John D. Collen for collecting the Woodside Creek samples, Robert N. Clayton for the use of his mass spectrometer, and Monica M. Grady for preliminary isotope data. This work was supported by NSF and NASA grants and by a NASA fellowship to W.S.W.

Received 17 February; accepted 13 June 1988.

1. Wolbach, W. S., Lewis, R. S. & Anders, E. *Science* **230**, 167–170 (1985).
2. Alvarez, L. W., Alvarez, W., Asaro, F. & Michel, H. V. *Science* **208**, 1095–1108 (1980).
3. Alvarez, W. *EOS* **67**, 649–658 (1986).
4. Hallam, A. *Science* **238**, 1237–1242 (1987).
5. Officer, C. B., Hallam, A., Drake, C. L. & Devine, J. D. *Nature* **326**, 143–149 (1987).
6. Alvarez, L. W. *Physics Today* **40**, 24–33 (1987).
7. Strong, C. P. *N.Z. J. Geol. Geophys.* **20**, 687–696 (1977).
8. Brooks, R. R. *et al. Science* **226**, 539–542 (1984).
9. Strong, C. P. *N.Z. geol. Survey Rec.* **3**, 47–51 (1985).
10. Brooks, R. R. *et al. Geology* **14**, 727–729 (1986).
11. Herring, J. R. in *The Carbon Cycle and Atmospheric CO<sub>2</sub>: Natural Variations Archean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 419–442 (American Geophysical Union, Washington, DC, 1985).
12. Tang, M. thesis, Univ. Chicago (1987).
13. Medalia, A. I. & Heckman, F. A. *Carbon* **7**, 567–582 (1969).
14. Anders, E., Wolbach, W. S. & Lewis, R. S. *Science* **234**, 261–264 (1986).
15. Officer, C. B. & Ekdale, A. A. *Science* **234**, 261–264 (1986).
16. Kastner, M., Asaro, F., Michel, H. V., Alvarez, W. & Alvarez, L. W. *Science* **226**, 137–143 (1984).
17. Argyle, E. *Science* **234**, 261–264 (1986).
18. Toon, O. B. *et al. Spec. Pap. geol. Soc. Am.* **190**, 187–200 (1982).
19. Strong, C. P. *et al. Geochim. cosmochim. Acta* **51**, 2769–2777 (1987).
20. Brooks, R. R. *et al. Geology* **13**, 738–740 (1985).
21. Smit, J. & Kyte, F. T. *Nature* **310**, 403–405 (1984).
22. Bohor, B. F., Modreski, P. J. & Foord, E. E. *Lun. planet. Sci.* **16**, 79–80 (1985).
23. Ledbetter, M. T. & Sparks, R. S. J. *Geology* **7**, 240–244 (1979).
24. Wilson, L. *Geophys. J. R. astr. Soc.* **30**, 381–392 (1972).
25. Preisinger, A. *et al. Nature* **322**, 794–799 (1986).

26. Kyte, F. T., Smit, J. & Wasson, J. T. *Earth planet. Sci. Lett.* **73**, 183–195 (1985).
27. Gilmour, I. & Anders, E. *Lunar planet. Sci.* **19**, 389–390 (1988).
28. Hansen, H. J. *et al. in Global Bio-Events* (ed. Walliser, O. H.) 381–384 (Springer, Berlin, 1986).
29. Prinn, R. G. & Fegley, B. Jr. *Earth planet. Sci. Lett.* **83**, 1–15 (1987).
30. Smith, B. N. *BioScience* **22**, 226–231 (1972).
31. Deines, P. in *Handbook of Environmental Isotope Geochemistry* Vol. 1 (eds Fritz, P. & Fortes, J. Ch.) 329–407 (Elsevier, Amsterdam, 1980).
32. Holland, H. D. *The Chemistry of the Atmosphere and Oceans* (Wiley, New York, 1978).
33. Tissot, B. P. & Welte, D. H. *Petroleum Formation and Occurrence*, 2nd edn (Springer, Berlin, 1984).
34. Gilmour, I., Boyd, S. R. & Pillinger, C. T. *Meteoritics* **23** (in the press).
35. Andreea, M. O. *Science* **220**, 1148–1151 (1983).
36. Griffin, J. J. & Goldberg, E. D. *Science* **206**, 563–565 (1979).
37. Simoneit, B. R. T. & Beller, H. R. *Init. Rep. DSDP 93*, Ch. 52 (eds van Hinte, J. E. *et al.*) 1211–1215 (US Govt. Printing Office, Washington, DC, 1987).
38. Ramdahl, T. *Nature* **306**, 580–582 (1983).
39. Seiler, W. & Crutzen, P. J. *Climatic Change* **2**, 207–247 (1980).
40. National Academy of Sciences *The Effects on the Atmosphere of a Major Nuclear Exchange* (National Academy, Washington, DC, 1985).
41. Pittock, A. B. *et al. Environmental Consequences of Nuclear War, Vol. 1: Physical and Atmospheric Effects* SCOPE 28 (Wiley, Chichester, 1986).
42. Berner, R. A. & Landis, G. P. *Science* **239**, 1406–1409 (1988).
43. Watson, A. J., Lovelock, J. E., & Margulis, L. *BioSystems* **10**, 293 (1978).
44. Emiliani, C., Kraus, E. B. & Shoemaker, E. M. *Earth planet. Sci. Lett.* **55**, 317–334 (1981).
45. Bohor, B. F., Modreski, P. J. & Foord, E. E. *Science* **236**, 705–709 (1987).
46. Cisowski, S. M. & Fuller, M. *Science* **234**, 261–264 (1986).
47. Lewis, J. S., Watkins, G. H., Hartman, H. & Prinn, R. G. *Spec. Pap. geol. Soc. Am.* **190**, 215–221 (1982).
48. Kroto, H. W. & McKay, K. *Nature* **331**, 328–331 (1988).

# Reversibility of acidification shown by whole-catchment experiments

R. F. Wright\*, E. Lotse† & A. Semb‡

\* Norwegian Institute for Water Research, Box 33 Blindern, 0313 Oslo, Norway

† Department of Soil Sciences, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden

‡ Norwegian Institute for Air Research, Box 64, 2001 Lillestrøm, Norway

*Manipulation experiments are being conducted in Norway to examine the effects of drastic changes in precipitation chemistry on soil and surface-water acidification. At a clean area in western Norway, two pristine catchments are being acidified by addition of H<sub>2</sub>SO<sub>4</sub> and H<sub>2</sub>SO<sub>4</sub> + HNO<sub>3</sub>, respectively. At an acidified catchment in southernmost Norway, ambient acid precipitation is excluded by means of a roof and clean precipitation added beneath.*

ACID deposition, acidification of surface waters and loss of fish populations occur over large regions of Europe and eastern North America<sup>1-3</sup>. International efforts to reduce the emissions of SO<sub>2</sub> and NO<sub>x</sub> are based in part on the premise that such reductions will restore acidified waters. To test this premise we are conducting whole-catchment manipulations in which the acid loading is changed experimentally. The RAIN project (Reversing Acidification In Norway) is composed of two experiments: artificial acidification of two pristine catchments in western Norway (Sogndal site) and exclusion of ambient acid deposition at an acidified catchment in southernmost Norway (Risdsalsheia site)<sup>4-7</sup>. We report here on the changes in soil and runoff chemistry after four years of treatment.

Acid addition at Sogndal has caused major changes in runoff chemistry (Table 1, Fig. 1), including short-term episodes of acid, aluminium-rich runoff and recovery between episodes. Sulphate concentrations have increased from ~20–25 µequiv. l<sup>-1</sup> to 50–57 µequiv. l<sup>-1</sup> at SOG2 (H<sub>2</sub>SO<sub>4</sub> addition) and 35–45 µequiv. l<sup>-1</sup> at SOG4 (H<sub>2</sub>SO<sub>4</sub> + HNO<sub>3</sub> addition). These levels are ~50% of expected steady-state concentrations at which sulphate flux out equals sulphate flux in. Addition of HNO<sub>3</sub> has caused only minor increases in nitrate in runoff. The increased sulphate concentrations have been about 50 per cent compensated by increased concentrations of base cations (mainly Ca<sup>2+</sup> and Mg<sup>2+</sup>) and about 50 per cent by decreased

alkalinity. Continued acid addition and loss of base cations at this rate could result in significant soil acidification within a few decades.

Acid exclusion at the acidified catchment in southernmost Norway has resulted in lower concentrations of the strong-acid anions NO<sub>3</sub><sup>-</sup> (from 35 to 7 µequiv. l<sup>-1</sup>) and SO<sub>4</sub><sup>2-</sup> (from 110 to 53 µequiv. l<sup>-1</sup>) in runoff relative to both the roofed control catchment and open catchment (Table 2, Fig. 2). The decline in strong-acid anion concentrations was compensated for partly by a decrease in concentrations of base cations (55%) and partly by an increase in alkalinity (45%). The input-output budgets indicate that the acid exclusion has reversed soil acidification at Risdsalsheia. The effect of organic acids on the pH of runoff has increased in importance as the experiment has proceeded. The acid-exclusion experiment shows that chemical changes caused by acid deposition are largely reversible.

## Sogndal: acid addition

The Sogndal site is a pristine but sensitive area in western Norway presently receiving only minor acidic precipitation (pH 4.8). The site is located 900 m above sea level and on gneissic bedrock, with patchy, thin (average depth 30 cm) and poorly developed soils (Lithic Haplumbrept) having pH (H<sub>2</sub>O) 4.5–5.5, and alpine vegetation<sup>8</sup>. Four catchments are being studied: catchment SOG2 (7,220 m<sup>2</sup>) receives H<sub>2</sub>SO<sub>4</sub>, catchment SOG4

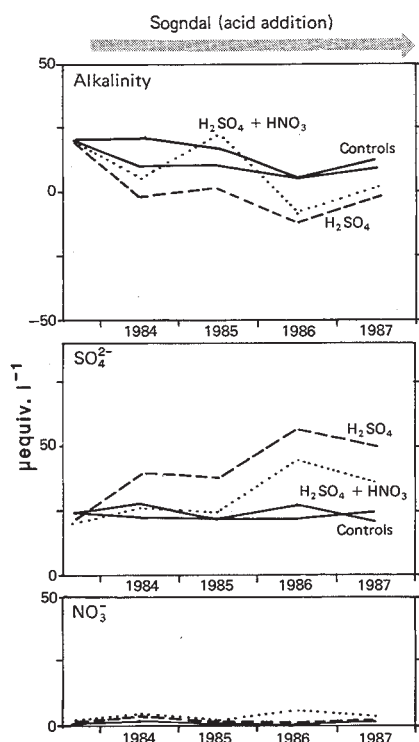
Table 1 Volume-weighted mean concentrations of major ions in precipitation and runoff at the Sogndal catchments

|                               | Precipitation<br>SOG1, SOG3 |       |      |      |      | Runoff<br>SOG1<br>control |       |      |      |      | Runoff<br>SOG3<br>control |       |      |      |      | Runoff<br>SOG2<br>H <sub>2</sub> SO <sub>4</sub> addition |       |      |      |      | Runoff<br>SOG4<br>H <sub>2</sub> SO <sub>4</sub> + HNO <sub>3</sub> addition |       |      |      |      |
|-------------------------------|-----------------------------|-------|------|------|------|---------------------------|-------|------|------|------|---------------------------|-------|------|------|------|-----------------------------------------------------------|-------|------|------|------|------------------------------------------------------------------------------|-------|------|------|------|
|                               | pre                         | 1984  | 1985 | 1986 | 1987 | pre                       | 1984  | 1985 | 1986 | 1987 | pre                       | 1984  | 1985 | 1986 | 1987 | pre                                                       | 1984  | 1985 | 1986 | 1987 | pre                                                                          | 1984  | 1985 | 1986 | 1987 |
| H <sub>2</sub> O              | 320                         | 1,126 | 928  | 957  | 925  | 320                       | 1,026 | 828  | 857  | 790  | 320                       | 1,026 | 828  | 857  | 790  | 320                                                       | 1,086 | 891  | 925  | 683  | 320                                                                          | 1,086 | 891  | 930  | 718  |
| H <sup>+</sup>                | 12                          | 20    | 15   | 11   | 18   | 1                         | 2     | 2    | 1    | 2    | 3                         | 4     | 1    | 1    | 2    | 3                                                         | 8     | 4    | 8    | 7    | 1                                                                            | 4     | 4    | 6    | 4    |
| pH                            | 4.9                         | 4.7   | 4.8  | 5.0  | 4.7  | 6.0                       | 5.7   | 5.7  | 6.0  | 5.8  | 5.5                       | 5.4   | 6.0  | 6.0  | 5.7  | 5.5                                                       | 5.1   | 5.4  | 5.1  | 5.1  | 6.0                                                                          | 5.4   | 5.4  | 5.2  | 5.4  |
| Na <sup>+</sup>               | 44                          | 72    | 32   | 22   | 28   | 38                        | 88    | 35   | 33   | 38   | 50                        | 74    | 33   | 30   | 38   | 53                                                        | 59    | 31   | 35   | 38   | 38                                                                           | 51    | 19   | 29   | 37   |
| K <sup>+</sup>                | 6                           | 6     | 4    | 6    | 6    | 0                         | 11    | 2    | 2    | 4    | 0                         | 2     | 1    | 0    | 1    | 3                                                         | 4     | 2    | 2    | 5    | 3                                                                            | 2     | 1    | 2    | 3    |
| Ca <sup>2+</sup>              | 6                           | 9     | 4    | 3    | 6    | 19                        | 20    | 16   | 18   | 20   | 19                        | 19    | 16   | 18   | 19   | 19                                                        | 22    | 22   | 28   | 29   | 22                                                                           | 26    | 20   | 27   | 26   |
| Mg <sup>2+</sup>              | 9                           | 16    | 4    | 5    | 5    | 9                         | 15    | 8    | 8    | 10   | 9                         | 17    | 8    | 8    | 11   | 12                                                        | 13    | 11   | 14   | 13   | 9                                                                            | 12    | 9    | 11   | 10   |
| Al <sup>3+</sup>              | —                           | —     | —    | —    | —    | 0                         | 1     | 0    | 0    | 0    | 0                         | 0     | 0    | 0    | 1    | 0                                                         | 3     | 2    | 2    | 3    | 0                                                                            | 1     | 0    | 1    | 3    |
| NH <sub>4</sub> <sup>+</sup>  | 3                           | 9     | 6    | 7    | 14   | 0                         | 6     | 5    | —    | —    | 0                         | 1     | 0    | —    | —    | 0                                                         | 3     | 0    | —    | —    | 0                                                                            | 1     | 1    | —    | —    |
| NO <sub>3</sub> <sup>-</sup>  | 3                           | 9     | 8    | 5    | 11   | 3                         | 2     | 1    | 1    | 1    | 0                         | 2     | 0    | 0    | 0    | 3                                                         | 4     | 1    | 1    | 2    | 0                                                                            | 4     | 1    | 6    | 3    |
| Cl <sup>-</sup>               | 47                          | 81    | 26   | 25   | 30   | 34                        | 89    | 26   | 28   | 37   | 41                        | 78    | 25   | 27   | 36   | 62                                                        | 60    | 26   | 34   | 36   | 38                                                                           | 57    | 26   | 26   | 35   |
| SO <sub>4</sub> <sup>2-</sup> | 16                          | 31    | 16   | 20   | 24   | 22                        | 28    | 22   | 27   | 21   | 22                        | 23    | 22   | 23   | 23   | 22                                                        | 40    | 38   | 57   | 50   | 19                                                                           | 26    | 24   | 45   | 36   |
| HCO <sub>3</sub> <sup>-</sup> | —                           | —     | —    | —    | —    | 9                         | 15    | 13   | 4    | 12   | 6                         | 1     | 5    | 4    | 8    | 9                                                         | 5     | 2    | 0    | 0    | 12                                                                           | 6     | 6    | 1    | 8    |
| Sum <sup>+</sup>              | 80                          | 132   | 65   | 54   | 77   | 67                        | 143   | 68   | 62   | 74   | 81                        | 117   | 59   | 57   | 72   | 90                                                        | 112   | 72   | 89   | 95   | 73                                                                           | 97    | 54   | 76   | 83   |
| Sum <sup>-</sup>              | 66                          | 126   | 50   | 48   | 65   | 68                        | 134   | 65   | 60   | 71   | 69                        | 104   | 52   | 54   | 75   | 96                                                        | 109   | 67   | 92   | 88   | 69                                                                           | 93    | 57   | 78   | 82   |

The periods refer to pre-treatment 1 September to 14 November 1983 (pre), 15 November 1983 to 12 November 1984 (1984), 13 November 1984 to 1 November 1985 (1985), 2 November 1985 to 12 November 1986 (1986) and 12 November 1986 to 15 November 1987 (1987). Treatment began in April 1984. Al is measured as reactive Al and non-labile monomeric Al (NLAL); labile monomeric Al is obtained by difference<sup>24</sup>; concentrations of ionic Al species are calculated by the procedure of Schecher and Driscoll<sup>13</sup>; values here are the sum of all positively charged species. Dash (—) indicates no data. Units: µequiv. l<sup>-1</sup>; H<sub>2</sub>O mm.

Added to SOG2: 1984, 60 mm, 70 mequiv. m<sup>-2</sup> H<sub>2</sub>SO<sub>4</sub>; 1985, 61 mm, 100 mequiv. m<sup>-2</sup> H<sub>2</sub>SO<sub>4</sub>; 1986, 68 mm, 100 mequiv. m<sup>-2</sup> H<sub>2</sub>SO<sub>4</sub>; 1987, 76 mm, 100 mequiv. m<sup>-2</sup> H<sub>2</sub>SO<sub>4</sub>. Added to SOG4: 1984 58 mm, 35 mequiv. m<sup>-2</sup> H<sub>2</sub>SO<sub>4</sub> + 35 mequiv. m<sup>-2</sup> HNO<sub>3</sub>; 1985 65 mm, 50 mequiv. m<sup>-2</sup> H<sub>2</sub>SO<sub>4</sub> + 50 mequiv. m<sup>-2</sup> HNO<sub>3</sub>; 1986 73 mm, 50 mequiv. m<sup>-2</sup> H<sub>2</sub>SO<sub>4</sub> + 50 mequiv. m<sup>-2</sup> HNO<sub>3</sub>; 1987 74 mm, 50 mequiv. m<sup>-2</sup> H<sub>2</sub>SO<sub>4</sub> + 50 mequiv. m<sup>-2</sup> HNO<sub>3</sub>.



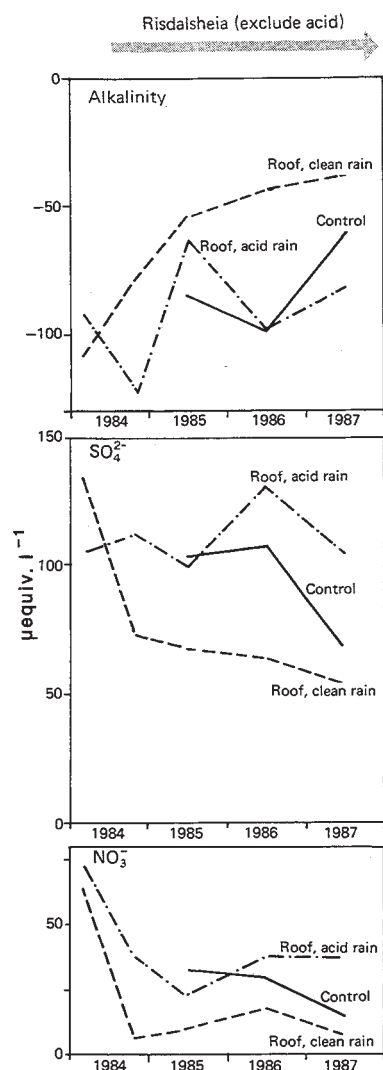


**Fig. 1** Alkalinity, sulphate and nitrate in runoff (volume-weighted annual mean concentrations) from the experimental catchments at Sogndal. Alkalinity is defined as the sum of charges of base cations less the sum of charges of strong-acid anions. Acid addition was 70 mequiv.  $\text{m}^{-2} \text{yr}^{-1}$  in 1984 and 100 mequiv.  $\text{m}^{-2} \text{yr}^{-1}$  in 1985–87 either as  $\text{H}_2\text{SO}_4$  or a 1:1 mixture of  $\text{H}_2\text{SO}_4 + \text{HNO}_3$ . See also Table 1.

(1,940  $\text{m}^2$ ) receives a 1:1 mixture of  $\text{H}_2\text{SO}_4 + \text{HNO}_3$ , and catchments SOG1 (96,300  $\text{m}^2$ ) and SOG3 (43,200  $\text{m}^2$ ) serve as untreated controls. The  $\text{H}_2\text{SO}_4$  treatment provides information on the effect of sulphuric acid alone; the  $\text{H}_2\text{SO}_4 + \text{HNO}_3$  treatment simulates a future acid deposition model in which  $\text{SO}_2$  emissions are reduced but  $\text{NO}_x$  emissions continue unabated. The treated catchments receive identical total loadings of  $\text{H}^+$ .

Acid addition, which began in April 1984, consists of application to the snowpack of 0.02 mm water at pH 1.9 and four or five events of 11 mm at pH 3.2 during the snow-free months. Acid is mixed with lake water from SOG1 and applied at 2  $\text{mm h}^{-1}$  by means of commercial irrigation equipment. Before and after each acid addition 2 mm of unacidified lake water is added to wet up and wash down the vegetation, respectively. The experiment is designed so that acid addition is obtained with a minimum of extra water (about 60 mm added to ~1,000 mm natural precipitation) added at an acidity that does not cause direct damage to vegetation. Total acid loading was 70 mequiv.  $\text{m}^{-2} \text{yr}^{-1}$  in 1984 and 100 mequiv.  $\text{m}^{-2} \text{yr}^{-1}$  in 1985, 1986 and 1987; these levels are typical of acidified areas in southernmost Norway.

Volume and chemical composition of natural precipitation at Sogndal are measured in weekly bulk samples collected at Haukås farm, 500 m above sea level and 3 km from the experimental catchments. In addition the water equivalent and chemical composition of the snowpack at the four catchments are measured in late winter each year. Because of the high relief in the area, precipitation volume differs substantially over short distances owing to snow-drifting and orographic effects. We assume that the snowcourse measurements give the best estimate of winter precipitation, and the gauge at Haukås farm the best estimate of summer precipitation.



**Fig. 2** Alkalinity, sulphate and nitrate in runoff (volume-weighted annual mean concentrations) from the experimental catchments at Risdalsheia. Alkalinity is defined as the sum of charges of base cations less the sum of charges of strong-acid anions. KIM catchment is covered by a roof and receives 'clean' precipitation. EGIL catchment is also roofed but receives recycled ambient acid precipitation. ROLF catchment is open and receives ambient acid precipitation. See also Table 2.

Discharge is gauged continuously by weir and level recorder at the outlet of catchments SOG1, SOG2 and SOG4. Discharge from SOG3 is not gauged and is assumed proportional to that at SOG1. Runoff samples for chemical analysis are collected at least weekly at the outlet of each catchment; sampling frequency is increased to 2–7 times weekly during snowmelt. Additional samples from SOG2 and SOG4 are collected every 2 hours during and immediately following acid additions, and daily for the five days after acid addition<sup>4–7</sup>.

Four years of acid addition at Sogndal have resulted in major changes in runoff chemistry (Table 1, Fig. 1). Each event of acid addition gives a short-term episode of acid, aluminium-rich runoff. Recovery between episodes was rapid and full in the first year but increasingly slower and less complete in subsequent years. Volume-weighted mean sulphate concentrations have increased from about 20–25  $\mu\text{equiv. l}^{-1}$  (pre-treatment and control catchments SOG1 and SOG3) to 50–57  $\mu\text{equiv. l}^{-1}$  at SOG2 and 35–45  $\mu\text{equiv. l}^{-1}$  at SOG4. These levels are ~50% of the

**Table 2** Volume-weighted mean concentrations of major ions in precipitation and runoff at the Risdalsheia catchments

|                               | EGIL catchment (roof, acid rain) |      |      |       |     |        |      |      |       |      | KIM catchment (roof, clean rain) |      |      |     |      |        |      |      |      |       | ROLF catchment (no roof, acid rain) |       |      |      |       |        |      |      |       |       |  |  |
|-------------------------------|----------------------------------|------|------|-------|-----|--------|------|------|-------|------|----------------------------------|------|------|-----|------|--------|------|------|------|-------|-------------------------------------|-------|------|------|-------|--------|------|------|-------|-------|--|--|
|                               | Deposition                       |      |      |       |     | Runoff |      |      |       |      | Deposition                       |      |      |     |      | Runoff |      |      |      |       | Deposition                          |       |      |      |       | Runoff |      |      |       |       |  |  |
|                               | 1984                             | 1985 | 1986 | 1987  | pre | 1984   | 1985 | 1986 | 1987  | 1984 | 1985                             | 1986 | 1987 | pre | 1984 | 1985   | 1986 | 1987 | 1984 | 1985  | 1986                                | 1987  | 1984 | 1985 | 1986  | 1987   | 1984 | 1985 | 1986  | 1987  |  |  |
| H <sub>2</sub> O              | 654                              | 692  | 771  | 1,134 | 417 | 547    | 545  | 570  | 1,103 | 506  | 596                              | 686  | 905  | 417 | 410  | 425    | 531  | 786  | 872  | 1,409 | 1,516                               | 1,524 | —    | 999  | 1,388 | 1,324  | —    | 999  | 1,388 | 1,324 |  |  |
| H <sup>+</sup>                | 58                               | 84   | 83   | 71    | 77  | 110    | 75   | 88   | 78    | 28   | 47                               | 39   | 30   | 89  | 95   | 72     | 72   | 77   | 53   | 72    | 77                                  | 63    | —    | 87   | 93    | 70     | —    | 87   | 93    | 70    |  |  |
| pH                            | 4.2                              | 4.1  | 4.1  | 4.2   | 4.1 | 4.0    | 4.1  | 4.1  | 4.1   | 4.6  | 4.3                              | 4.4  | 4.5  | 4.1 | 4.0  | 4.1    | 4.1  | 4.1  | 4.3  | 4.1   | 4.1                                 | 4.2   | —    | 4.1  | 4.0   | 4.2    | —    | 4.1  | 4.0   | 4.2   |  |  |
| Na <sup>+</sup>               | 78                               | 68   | 62   | 96    | 101 | 91     | 75   | 93   | 96    | 75   | 92                               | 96   | 71   | 134 | 83   | 99     | 89   | 86   | 73   | 56    | 102                                 | 60    | —    | 84   | 101   | 85     | —    | 84   | 101   | 85    |  |  |
| K <sup>+</sup>                | 5                                | 4    | 4    | 4     | 10  | 5      | 7    | 10   | 8     | 0    | 0                                | 3    | 3    | 10  | 5    | 12     | 8    | 7    | 5    | 5     | 5                                   | 3     | —    | 4    | 4     | 3      | —    | 4    | 4     | 3     |  |  |
| Ca <sup>2+</sup>              | 9                                | 9    | 9    | 10    | 14  | 15     | 15   | 19   | 19    | 4    | 2                                | 6    | 4    | 19  | 12   | 14     | 13   | 11   | 9    | 8     | 12                                  | 8     | —    | 15   | 20    | 15     | —    | 15   | 20    | 15    |  |  |
| Mg <sup>2+</sup>              | 17                               | 16   | 18   | 22    | 29  | 24     | 20   | 28   | 25    | 16   | 22                               | 20   | 16   | 31  | 15   | 19     | 19   | 18   | 16   | 11    | 25                                  | 17    | —    | 25   | 30    | 21     | —    | 25   | 30    | 21    |  |  |
| Al <sup>3+</sup>              | —                                | —    | —    | —     | 24  | 20     | 10   | 15   | 14    | —    | —                                | —    | —    | 17  | 5    | 8      | 9    | 10   | —    | —     | —                                   | —     | —    | 8    | 12    | 12     | —    | 8    | 12    | 12    |  |  |
| NH <sub>4</sub> <sup>+</sup>  | 43                               | 53   | 39   | 37    | 24  | 9      | 22   | 33   | 26    | 4    | 8                                | 6    | 8    | 17  | 2    | 12     | 13   | 6    | 41   | 44    | 42                                  | 36    | —    | 14   | 14    | 6      | —    | 14   | 14    | 6     |  |  |
| NO <sub>3</sub> <sup>-</sup>  | 47                               | 53   | 43   | 46    | 72  | 37     | 22   | 38   | 37    | 14   | 20                               | 9    | 12   | 62  | 5    | 9      | 17   | 7    | 44   | 50    | 49                                  | 40    | —    | 31   | 29    | 14     | —    | 31   | 29    | 14    |  |  |
| Cl <sup>-</sup>               | 84                               | 75   | 83   | 101   | 91  | 117    | 81   | 111  | 103   | 89   | 111                              | 105  | 84   | 120 | 117  | 132    | 104  | 106  | 79   | 56    | 119                                 | 69    | —    | 92   | 130   | 108    | —    | 92   | 130   | 108   |  |  |
| SO <sub>4</sub> <sup>2-</sup> | 81                               | 85   | 82   | 73    | 106 | 112    | 99   | 131  | 106   | 26   | 37                               | 35   | 25   | 137 | 73   | 68     | 64   | 53   | 79   | 77    | 87                                  | 66    | —    | 104  | 108   | 68     | —    | 104  | 108   | 68    |  |  |
| A <sup>2-</sup>               | —                                | —    | —    | —     | 10  | 12     | 24   | 11   | 20    | —    | —                                | —    | —    | 0   | 22   | 29     | 38   | 49   | —    | —     | —                                   | —     | —    | 14   | 6     | 22     | —    | 14   | 6     | 22    |  |  |
| Sum <sup>+</sup>              | 210                              | 234  | 215  | 240   | 279 | 276    | 225  | 288  | 266   | 127  | 171                              | 171  | 132  | 317 | 217  | 237    | 223  | 215  | 197  | 196   | 263                                 | 187   | —    | 238  | 274   | 212    | —    | 238  | 274   | 212   |  |  |
| Sum <sup>-</sup>              | 212                              | 213  | 208  | 220   | 279 | 276    | 225  | 288  | 266   | 129  | 168                              | 149  | 121  | 319 | 217  | 237    | 223  | 215  | 202  | 183   | 255                                 | 175   | —    | 240  | 273   | 212    | —    | 240  | 273   | 212   |  |  |
| NLAL                          |                                  |      |      |       |     | 101    | 188  | 231  | 191   | 112  |                                  |      |      |     |      | 149    | 280  | 289  | 233  | 185   |                                     |       |      |      |       | —      | 157  | 125  | 97    |       |  |  |
| TOC                           |                                  |      |      |       |     | 5.3    | 9.1  | 14.1 | 10.7  | 9.4  |                                  |      |      |     |      | 8.4    | 17.3 | 19.8 | 15.6 | 15.4  |                                     |       |      |      |       | —      | 12.9 | 10.1 | 9.7   |       |  |  |
| c.d.                          |                                  |      |      |       |     | 1.9    | 1.3  | 1.7  | 1.0   | 1.2  |                                  |      |      |     |      | 0      | 1.3  | 1.5  | 2.4  | 2.5   |                                     |       |      |      |       | —      | 0.9  | 0.8  | 2.3   |       |  |  |
| Alk.                          |                                  |      |      |       |     | -91    | -122 | -63  | -97   | -72  |                                  |      |      |     |      | -108   | -78  | -53  | -43  | -38   |                                     |       |      |      |       | —      | -85  | -98  | -60   |       |  |  |

Precipitation data for 1984 are for  $\frac{1}{2}$  yr only and begin 13 June. The periods refer to pre-treatment 25 March to 12 June 1984 (pre), 13 June to 13 December 1984 (1984), 14 December 1984 to 14 November 1985 (1985), 15 November 1985 to 18 December 1986 (1986), and 19 December 1986 to 23 November 1987 (1987). Organic anions (A<sup>-</sup>) are determined by difference from the ionic balance. Organic acid charge density (c.d. in  $\mu\text{equiv. per mg C}$ ) is defined as A<sup>-</sup> divided by TOC. Deposition includes wet and dry with concentrations obtained by dividing total flux by mm of precipitation. Alkalinity (Alk) is defined as the sum of base cations (Na<sup>+</sup>, K<sup>+</sup>, Ca<sup>2+</sup>, Mg<sup>2+</sup> and NH<sub>4</sub><sup>+</sup>) less the sum of strong-acid anions (NO<sub>3</sub><sup>-</sup>, Cl<sup>-</sup> and SO<sub>4</sub><sup>2-</sup>). Dry deposition is estimated for sea salts from chloride budget at EGIL catchment, for sulphate and nitrate from concentrations of SO<sub>2</sub> gas, SO<sub>4</sub><sup>2-</sup> particulates and NO<sub>2</sub> gas measured daily at the station Birkenes, 5 km west of Risdalsheia<sup>35</sup>, NH<sub>4</sub><sup>+</sup> from dry deposit of SO<sub>4</sub><sup>2-</sup> particulates and a molar ratio of 1.5, and H<sup>+</sup> by ionic balance<sup>35</sup>. Runoff data from 1984 for ROLF catchment are incomplete and not included. Deposition data for ROLF catchment in 1984–85 are daily samples from Birkenes<sup>35</sup>. Dash (—) indicates no data. Units  $\mu\text{equiv. l}^{-1}$ , H<sub>2</sub>O mm yr<sup>-1</sup>, NLAL  $\mu\text{g Al per litre}$ , TOC mg C per litre, charge density  $\mu\text{equiv. per mg C}$ . See also text to Table 1.

**Table 3** Pools of total sulphur and readily available sulphate in soil and sulphate fluxes at the experimental catchments at Sogndal (acid addition) and Risdalsheia (acid exclusion)

| Sogndal                                                  |             |             |             |          |          | Risdalsheia                                              |             |         |          |          |          |
|----------------------------------------------------------|-------------|-------------|-------------|----------|----------|----------------------------------------------------------|-------------|---------|----------|----------|----------|
| SOG1 (control)                                           |             |             |             |          |          | KIM (roof, clean rain)                                   |             |         |          |          |          |
|                                                          | 1984        | 1986        | Δ86–84      | <i>n</i> | <i>p</i> |                                                          | 1983        | 1986    | Δ86–83   | <i>n</i> | <i>p</i> |
| Soil pools                                               |             |             |             |          |          | Soil pools                                               |             |         |          |          |          |
| Total S                                                  | 4,500±500   | 3,800±800   | –700±800    | 4        | n.s.     | Total S                                                  | 1,700±1,000 | 910±450 | –800±450 | 9        | 0.01     |
| Water-soluble<br>+adsorbed SO <sub>4</sub> <sup>2–</sup> | 370±390     | 220±80      | –150±420    | 4        | n.s.     | Water-soluble<br>+adsorbed SO <sub>4</sub> <sup>2–</sup> | 82±36       | 57±28   | –35±30   | 9        | 0.01     |
| Fluxes 1984–86: in 69, out 70, Δin–out                   | –1          |             |             |          |          | Fluxes 1984–86: in 59, out 93, Δin–out                   | –34         |         |          |          |          |
| SOG2 (H <sub>2</sub> SO <sub>4</sub> )                   |             |             |             |          |          | EGIL (roof, acid rain)                                   |             |         |          |          |          |
|                                                          | 1984        | 1986        | Δ86–84      | <i>n</i> | <i>p</i> |                                                          | 1983        | 1986    | Δ86–83   | <i>n</i> | <i>p</i> |
| Soil pools                                               |             |             |             |          |          | Soil pools                                               |             |         |          |          |          |
| Total S                                                  | 2,660±1,130 | 3,790±1,990 | 1,200±1,640 | 14       | 0.025    | Total S                                                  | 1,070±70    | 880±210 | –190±100 | 6        | 0.01     |
| Water-soluble<br>+adsorbed SO <sub>4</sub> <sup>2–</sup> | 420±310     | 860±960     | 440±720     | 14       | 0.05     | Water-soluble<br>+adsorbed SO <sub>4</sub> <sup>2–</sup> | 40±15       | 46±12   | 6±14     | 6        | n.s.     |
| Fluxes 1984–86: in 339, out 130, Δin–out                 | 209         |             |             |          |          | Fluxes 1984–86: in 175, out 190, Δin–out                 | –15         |         |          |          |          |
| SOG4 (H <sub>2</sub> SO <sub>4</sub> +HNO <sub>3</sub> ) |             |             |             |          |          |                                                          |             |         |          |          |          |
|                                                          | 1984        | 1986        | Δ86–84      | <i>n</i> | <i>p</i> |                                                          |             |         |          |          |          |
| Soil pools                                               |             |             |             |          |          |                                                          |             |         |          |          |          |
| Total S                                                  | 4,000±1,000 | 4,400±1,200 | 400±1,400   | 4        | n.s.     |                                                          |             |         |          |          |          |
| Water-soluble<br>+adsorbed SO <sub>4</sub> <sup>2–</sup> | 440±300     | 380±320     | –50±180     | 4        | n.s.     |                                                          |             |         |          |          |          |
| Fluxes 1984–86: in 204, out 91, Δin–out                  | 113         |             |             |          |          |                                                          |             |         |          |          |          |

Soil data are from samples collected early autumn 1983 (Risdalsheia), 1984 (Sogndal) and 1986. Readily available sulphate taken as sum of water-soluble and adsorbed sulphate. Samples collected at 15-cm depth intervals and weighted for soil bulk-density and volume. Number of samples given by *n*; *p* refers to statistical significance of paired *t*-test of change in pool size. Treatment began in 1984 at Sogndal and mid-1984 at Risdalsheia. Units: mequiv. SO<sub>4</sub><sup>2-</sup> per m<sup>2</sup>.

expected steady-state concentrations at which sulphate flux in runoff equals sulphate deposition (wet plus dry), a characteristic of the untreated control catchments SOG1 and SOG3 (Table 2) and catchments in glaciated regions receiving acid deposition in Europe and North America<sup>9–12</sup>.

Addition of HNO<sub>3</sub>, on the other hand, has caused only minor increases in nitrate in runoff at SOG4. Although a short-term pulse of nitrate occurs during each acid addition, mass budget

calculations indicate that >90% of the added nitrate is retained in the catchment. The efficient retention of nitrate is probably due to uptake by the terrestrial vegetation; nitrogen is the growth-limiting nutrient for these ecosystems.

About 80% of the sulphate added to catchment SOG2 over the four-year period 1984–87 has been retained and apparently stored in the soil. This addition is sufficient to increase the pool of readily available sulphate (adsorbed and water-soluble sul-



phate) by  $\sim 70\%$  from the  $420 \pm 310$  mequiv.  $m^{-2}$  measured in August 1984 (Table 3). Resampling in August 1986 indicated that the pool of readily available sulphate in the soil had increased by  $\sim 440 \pm 720$  mequiv.  $m^{-2}$ ; the increase is significant at the 95% level. The relatively large uncertainties in these estimates of pool size are due to spatial heterogeneity in soil chemistry within the catchment. At catchment SOG4  $\sim 85\%$  of the  $185 \mu\text{equiv. } m^{-2}$  added has been retained; this amounts to a 35% increase in readily available sulphate (Table 3). Resampling indicated no significant change in sulphate pool. As the treatments continue, future measurements of sulphur fractions in the soils will reveal whether the retained sulphate is stored in the readily available pool or transformed to other sulphur forms. At both catchments the change in pool size inferred from the input-output budgets and the increase in sulphate concentrations in runoff over the four-year treatment period are consistent with the assumption that sulphate adsorption is the major soil-chemical process which accounts for the year-to-year trends in sulphate concentration in runoff.

The increased sulphate concentrations in runoff at both catchments are compensated  $\sim 50\%$  by increased concentrations of base cations (mainly Ca and Mg) and  $\sim 50\%$  by decreased alkalinity (defined as the sum of charges of base cations minus the sum of charges of strong-acid anions) (Table 1). By the third year of acid addition alkalinity was chronically negative at both SOG2 and SOG4; bicarbonate levels were negligible, and labile monomeric Al was of increasing importance. A toxicity experiment conducted in August 1986 showed that 2–4 days exposure to runoff from SOG2 was lethal to  $O^+$  salmon fry (S. Andersen, personal communication).

The concentrations of inorganic monomeric labile aluminium ( $Al_i$ ) in runoff increased dramatically with decreasing pH. Only minor amounts of organic monomeric aluminium ( $<20 \mu\text{g } Al_i l^{-1}$ ) are present, because concentrations of total organic carbon (TOC) are low ( $1\text{--}2 \text{ mg } C l^{-1}$ ). Aluminium solubility is not explained by equilibrium with respect to a single solid phase of aluminium hydroxide. In 1984, the first year of treatment, the relationship  $\log [Al_i] = \log [K] - b(pH)$  had an intercept  $\log [K] = 9.0$  and slope of 3, as would be expected for waters in equilibrium with natural gibbsite<sup>13</sup> (Fig. 3). In 1985 and 1986, however, Al concentrations were markedly lower at the same pH, and the slope was  $<3$  (Fig. 3). Then in 1987 Al levels were again higher, and the relationship was similar to that in 1984.

The action of several independent processes in the soils and stream may explain the Al data. The high levels of relatively soluble Al observed during the first acid additions in 1984 may reflect mobilization and release of Al from the streambed. Circum-neutral streams such as those at Sogndal contain a substantial store of Al which is rapidly released from the stream bed during acid episodes<sup>14–16</sup>. This aluminium is believed to originate in soil solution containing elevated levels of  $CO_2$ . As the soil solution enters the stream,  $CO_2$  degasses, the pH increases, and aluminium hydroxide precipitates<sup>17</sup>. Acid additions during the first year at Sogndal caused pulses of strong acid which then quickly redissolved this aluminium. Sulphuric acid added directly to a small stream at an adjacent catchment at Sogndal caused release of high concentrations of inorganic monomeric aluminium (S.A. Norton, personal communication). The  $\log [Al_i]$ -pH relationship had an intercept of 9.0 and a slope of 3.0, very similar to the 1984 data (Fig. 3).

The data from 1985 and 1986 indicate that in this stream bed the store of inorganic Al was depleted. Al concentrations in runoff probably reflected more complex relationships in the catchment soils. Several studies of Al in soil extracts have reported  $\log [Al_i]$ -pH relationships with slopes  $<3$  (refs 18, 19). The return to higher Al levels and higher apparent Al solubility in 1987 is more difficult to explain, but may indicate a resupply of stream-bed Al during the rather dry and warm summer.

Most of the increased flux of base cations in runoff from the

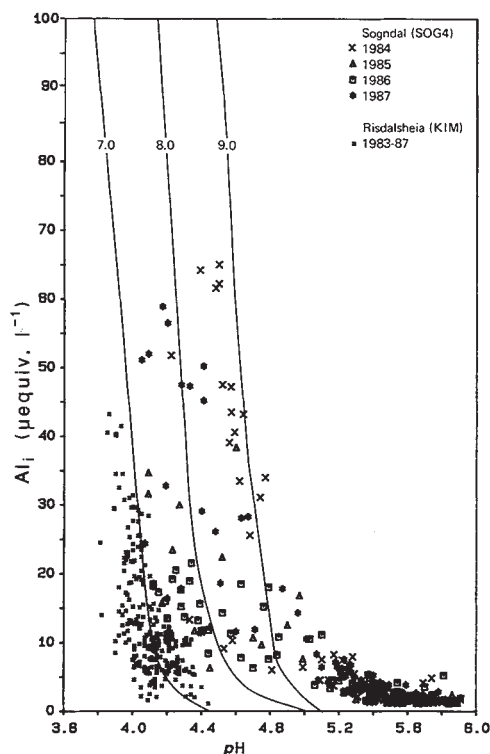


Fig. 3 Concentrations of  $Al_i$  and pH in runoff at catchments SOG4 ( $H_2SO_4 + HNO_3$  addition) at Sogndal and KIM (acid exclusion) at Risdalsheia. Labile monomeric Al was determined by the difference between reactive monomeric Al and non-labile monomeric Al measured by the procedure of Røgeberg and Henriksen<sup>35</sup>. Concentrations of  $Al_i$  were determined by the procedure of Schecher and Driscoll<sup>13</sup>. Also shown are curves of  $\log [Al_i] = \log [K] - 3pH$  for  $\log [K] = 7.0, 8.0$  and  $9.0$ . Natural gibbsite has  $\log [K]$  of 9.1 at  $25^\circ C$ .

acidified catchments at Sogndal is due to  $Ca^{2+}$  and  $Mg^{2+}$  (Table 1). Input-output budgets at the pristine control catchments can provide an estimate of the chemical weathering rate of base cations given the assumption that the catchments are at steady state with no significant change in cations stored in living or dead biomass or in the soil. The four-year data indicate that base cations are released by weathering at  $\sim 15$  mequiv.  $m^{-2} yr^{-1}$ ,  $10$  mequiv.  $m^{-2} yr^{-1}$  of which is  $Ca^{2+}$  and  $Mg^{2+}$ . The weathering estimate for  $Na^+$  is less certain because of the high natural flux of seasalt  $Na^+$ . The weathering rate at Sogndal is low relative to that reported from other ecosystems with similar soils and bedrock<sup>20</sup>.

The net flux of  $Ca^{2+}$  and  $Mg^{2+}$  from the treated catchments SOG2 and SOG4 is about 50% higher than that of the control catchments. This may be due either to depletion of the store of exchangeable base cations in the soil or to increased chemical weathering. Because added flux of base cations comes with the sulphate pulses after acid-addition events, cation exchange is probably the dominant process. We have estimated the size of this pool in 1984 at catchments SOG1, SOG2 and SOG4 by soil surveys and analysis of representative soil samples. The increased leaching of  $Ca^{2+}$  and  $Mg^{2+}$  during the four years of treatment corresponds to 1.7% of the exchange pool at SOG2 and 2.0% at SOG4. A significant decrease in soil base saturation by this process may thus occur after 1–3 decades of acid deposition. Such long-term soil acidification has been reported from Sweden<sup>21,22</sup> and Germany<sup>23</sup>.

Results obtained during the first four years of treatment indicate that these pristine, acid-sensitive catchments respond

rapidly to increases of acid deposition. Indeed, a single severe episode of acid precipitation may be sufficient to acidify runoff to the point at which the low- $pH$ , high- $Al$  water is toxic to fish. These catchments are typical of large regions of southern Norway. Over half of Norway is characterized by barren rock and soils <70 cm in depth<sup>1</sup>. The chronic acidification of runoff at Sogndal during the third and fourth years of treatment is caused entirely by an increase in acid deposition to levels typical for southernmost Norway. No land-use change or any other environmental change is responsible for the acidification at Sogndal.

### Risdalsheia: acid exclusion

The Risdalsheia site is located in a sensitive area in southernmost Norway currently receiving a high loading of acid deposition (precipitation has  $pH$  4.2; sulphate loading wet and dry is  $100 \text{ mequiv. m}^{-2} \text{ yr}^{-1}$ ). The catchments are situated 300 m above sea level and are characterized by exposed granitic bedrock (30–50% of surface) and thin, organic-rich, truncated podzolic soils (average depth 10–15 cm, maximum depth 50 cm) with  $pH(H_2O)$  3.9–4.5, and sparse cover of pine (*pinus sylvestris* L.) and birch (*Betula pubescens* L.)<sup>24</sup>. Acid exclusion is accomplished by means of a 1,200- $m^2$  transparent roof that completely covers the 860- $m^2$  KIM catchment. Incoming precipitation is collected from the roof and pumped through a filter and ion-exchange system. Seawater salts are added back at ambient levels and the clean precipitation is automatically applied beneath the roof above the canopy by means of a sprinkler system at the rate of  $2 \text{ mm h}^{-1}$ . An adjacent 400- $m^2$  catchment (EGIL) has also been covered with a roof and receives ambient acid precipitation by means of an identical sprinkling system. During the winter, artificial snow is added beneath the roofs each year by means of commercial snow-making equipment. Water from a nearby pond is ion-exchanged and sea salts added to make clean snow at KIM. At EGIL in 1985 sea salts and sulphuric acid were added to make acid snow; in 1986 and 1987 ambient acid snow was blown under the roof by means of a snowblower. The water equivalents of snow added were 115 mm in 1985 (KIM and EGIL), 180 mm (KIM) and 110 mm (EGIL) in 1986, and 150 mm (KIM) and 250 mm (EGIL) in 1987. A third uncovered catchment (ROLF) serves as reference. Treatment began in June 1984.

Precipitation volume and chemical composition are measured in weekly bulk samples collected on site. The volume of water applied to the two roofed catchments is measured continuously by meters installed in the sprinkler systems. All runoff from the three catchments is collected at fibreglass weirs and conducted in large-dimension hoses to 500-l tanks. The tanks are fitted with automatic valve and data-logging systems to record continuously the total volume of water leaving each catchment. Samples for chemical analysis are collected from each 10–20 mm of runoff (minimum once weekly) by automatic samplers. More frequent samples are collected after dry periods<sup>4–7</sup>.

Acid exclusion at the KIM catchment has resulted in lower concentrations of the strong-acid anions  $NO_3^-$  and  $SO_4^{2-}$  relative to both the roofed control catchment EGIL and the open ROLF catchment (Table 2, Fig. 2). Nitrate concentrations decreased by 60% within two weeks after onset of treatment. (That this decrease is due to the exclusion of nitrate and sulphate inputs in precipitation was demonstrated by a 1-week period in October 1984 during which a technical failure of the ion-exchange columns resulted in application of 50 mm water with nitrate at  $50 \text{ } \mu\text{equiv. l}^{-1}$ . Nitrate concentrations in runoff jumped immediately from 2 to  $25 \text{ } \mu\text{equiv. l}^{-1}$  and then declined to  $3\text{--}4 \text{ } \mu\text{equiv. l}^{-1}$  within 3 weeks<sup>4</sup>). Sulphate concentrations showed a general decline beginning about 4 months after onset of treatment and after 3.5 yr of treatment these concentrations are about 50% of that at EGIL (Table 2, Fig. 2).

At the KIM catchment, sulphate output continues to exceed inputs in wet and dry deposition. Net loss of sulphate during

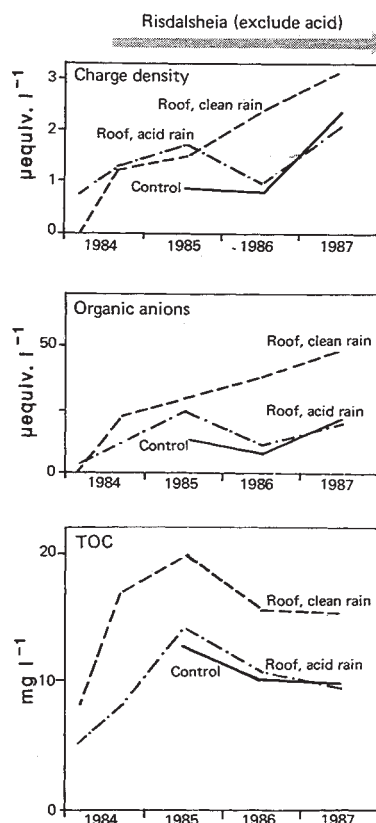


Fig. 4 Volume-weighted concentrations of charge density, organic anions ( $A^-$ ) and TOC in runoff from the experimental catchments at Risdalsheia. Organic anion concentrations are determined by the difference from the ionic balance; charge density is defined as  $A^-/\text{TOC}$ .

the first 3.5 years has been  $53 \text{ mequiv. m}^{-2}$ . This is equivalent to about 45% of the pool of readily available sulphate in the soil before treatment (Table 3).

The decline in strong-acid anion concentrations is compensated for partly by a decrease in concentrations of base cations (55%) and partly by an increase in alkalinity (45%). Changes in the concentrations of base cations are masked in part by year-to-year and catchment-to-catchment differences in the inputs of seawater salts. We estimate the seasalt component from Cl concentrations and the ionic composition of sea water. The non-marine fraction of base cations has decreased to 0–15  $\mu\text{equiv. l}^{-1}$  during the 3.5 yr of treatment, and was negative during the second half of 1984. Negative values imply that a portion of the seasalt base cations were retained in the catchment soils. Concentrations of ammonium in runoff are also lower than at the untreated control catchment (Table 2).

At Risdalsheia the concentrations of labile monomeric Al at a given  $pH$  are much lower than at Sogndal, and the  $\log [Al_i] - pH$  relationship approximately follows a line with intercept  $\log [K] = 6.8$  and slope of 3 (Fig. 3). The very low  $pH$  levels in runoff at Risdalsheia preclude a major  $pH$  change due to degassing of  $CO_2$ , and thus little Al should precipitate out in the stream bed.  $Al_i$  in the runoff at Risdalsheia probably reflects cation exchange in the soils. Concentrations of inorganic monomeric Al are negligible at  $pH > 4.4$ .

The input-output budgets indicate that relative to the control catchments the acid-exclusion catchment is retaining a portion of the base cations deposited in precipitation and released from weathering; the pool of exchangeable base cations is being replenished. The magnitude of this replenishment depends upon estimates for weathering rates in these catchments. Weathering rates cannot be obtained directly from input-output budgets at



the control catchments because a fraction of the base cations in runoff may come from the pool of exchangeable cations on the soil.

As the acid-exclusion experiment has proceeded, organic acids have become increasingly important. We estimate the concentrations of organic anions by difference from the ionic balance; uncertainty is  $\sim \pm 10 \mu\text{equiv. l}^{-1}$ . At the acid-exclusion catchment the concentration of organic anions has increased from  $\sim 22 \mu\text{equiv. l}^{-1}$  in 1984 to  $49 \mu\text{equiv. l}^{-1}$  in 1987 (Table 2, Fig. 4). This increase is due to increased dissociation of the organic acids and not to a change in concentration of total organic carbon (TOC); the charge density has increased from about  $1.3 \mu\text{equiv. per mg C}$  in 1984 to  $2.5$  in 1987 (Fig. 4). The organic carbon apparently has a maximum charge density of  $\sim 4.5 \mu\text{equiv. per mg C}$  and  $pK$  of  $\sim 4.2$  (ref. 7), similar to that reported from podzolic soils in New York, USA<sup>25</sup>. When fully dissociated,  $15 \text{ mg C per litre}$  will thus contribute  $\sim 70 \mu\text{equiv. l}^{-1}$  of organic anions. As the concentration of non-marine strong acid anions ( $\text{SO}_4^{2-} + \text{NO}_3^-$ ) declines to a new steady state of  $\sim 25 \mu\text{equiv. l}^{-1}$ , and with non-marine base cation concentrations of  $\sim 20 \mu\text{equiv. l}^{-1}$ , this  $pK$  and level of TOC will result in a  $pH$  of  $\sim 4.6$  in runoff.

At Risdalsheia the acid-exclusion experiment indicates that chemical changes caused by acid deposition are largely reversible. The mobile strong-acid anions  $\text{SO}_4^{2-}$  and  $\text{NO}_3^-$  in runoff responded rapidly to changes in input. At the KIM catchment this decrease in mobile anions has resulted in a reduction of strong acidity and a decrease in concentrations of base cations. So far only a small increase in  $pH$  of runoff has occurred because of buffering by organic acids.

Runoff from the treated catchment KIM contains high concentrations of organic carbon. This catchment is of necessity very small and thus has extremely thin and highly organic soils. Streams and lakes in southern Norway draining catchments with thicker, more mineral-rich soils generally have much lower levels of TOC, and organic anions play a minor role relative to sulphate and nitrate<sup>26</sup>. In Norway about 90% of lakes have TOC levels below  $6 \text{ mg C per litre}$  (ref. 27). In addition, base cation concentrations are generally higher owing to higher rates of chemical weathering.

Our data from the acid-exclusion experiment at Risdalsheia show that a decrease in acid deposition results in decreased concentrations of strong-acid anions in runoff. It is therefore likely that a major reduction in the flux of strong-acid anions will cause major change in  $pH$  and inorganic aluminium in clearwater lakes and streams in southern Norway; the water will be less toxic to fish and other aquatic organisms.

## Conclusion

At both Sogndal and Risdalsheia runoff from the treated catchments now differs substantially in chemical composition from that of the control catchments. New steady-state concentrations have not yet been reached. Runoff at Sogndal has been acidified to levels toxic to fish, and runoff at Risdalsheia has begun to recover to pre-acidification chemical composition. Acid addition at Sogndal has caused soil acidification; acid exclusion at Risdalsheia has initiated reversal of ongoing soil acidification.

The changes in runoff chemistry observed at the RAIN project sites can be explained qualitatively by the interaction of key soil-chemical processes such as sulphate adsorption, cation exchange, dissolution of  $\text{CO}_2$ , mobilization of aluminium and buffering by organic acids. These processes are central in understanding soil and water acidification by acid deposition<sup>26,28,29</sup>. A process-oriented model of soil and water acidification (MAGIC) applied to the first two years' data generally predicted the observed changes<sup>30</sup>. The RAIN project provides a unique data set for the calibration and evaluation of such predictive models. These models can, in turn, serve as the basis for extrapolation of the RAIN results to regional water and soil chemistry.

The RAIN project shows that within a few years changes in acid deposition cause major changes in surface water chemistry at sensitive sites. These sites are typical of large areas of southern Norway. The reversibility of acidification demonstrated here agrees with empirical data from Canada, the United States and Scotland<sup>3,31-34</sup>. As the RAIN project continues, the treated catchments will approach a new 'steady state' and provide information as to whether acidification of soil and runoff chemistry is fully reversible.

We thank S. Andersen, N. Christophersen, H. Efraimsson, E. Gjessing, R. Høgberget, T. Mindrebø, A. Rogne, H. M. Seip, B. Slettaune and R. Storhaug for assistance; N. Knagenhjelm, Sogn Televerk and Arendal Televerk, N. Dalaker, H. Haukås and A. Risdal for logistic support; and D. Brakke, B. Cosby, T. Frogner, M. Hauhs, A. Henriksen, S. Norton, R. Parnell and J. Reuss for helpful criticism. The RAIN project has received financial support from the Norwegian Institute for Water Research, the Norwegian Institute for Air Research, the Norwegian Ministry of Environment, The Royal Norwegian Council for Scientific and Industrial Research, the Ontario Ministry of the Environment, Environment Canada, the Swedish National Environmental Protection Board, and the Surface Water Acidification Programme (SWAP) (The Royal Society, the Norwegian Academy of Science and Letters, and the Royal Swedish Academy of Sciences).

Received 20 April; accepted 24 May 1988.

- Overrein, L. N., Seip, H. M. & Tollan, A. *Acid Precitation—Effects on Forest and Fish* (Rep. FR 19/80, SNSF-project, 1432 Ås, Norway, 1980).
- Wright, R. F. *Water Qual. Bull.* **8**, 137–142 (1983).
- Schindler, D. W. *Science* **239**, 149–157 (1988).
- Wright, R. F. *RAIN project. Annual report for 1984* (Acid Rain Res. Rep. 7/85, Norwegian Inst. Water Res., Oslo, 1985).
- Wright, R. F. & Gjessing, E. *RAIN project. Annual report for 1985* (Acid Rain Res. Rep. 9/86, Norwegian Inst. Water Res., Oslo, 1986).
- Wright, R. F. *et al.* *Water Air Soil Pollut.* **30**, 47–64 (1986).
- Wright, R. F. *RAIN project. Annual report for 1986* (Acid Rain Res. Rep. 13/87, Norwegian Inst. Water Res., Oslo, 1987).
- Lotse, E. & Otabbong, E. *Physicochemical properties of soils at Risdalsheia and Sogndal: RAIN project* (Acid Rain Res. Rep. 8/85, Norwegian Inst. Water Res., Oslo, 1985).
- Wright, R. F. & Johannessen, M. in *Ecological Impact of Acid Precipitation* (eds Drablos, D. & Tollan, A.) 250–251 (SNSF-project, 1432 Ås, Norway, 1980).
- Wright, R. F. *Predicting Acidification of North American Lakes* (Acid Rain Res. Rep. 4/83, Norwegian Inst. Water Res., Oslo, 1983).
- Kramer, J. R. *et al.* in *Acid Deposition Long-Term Trends*, 231–299 (National Academy of Sciences, Washington, DC, 1986).
- Rochelle, B. P., Church, M. R. & David, M. B. *Water Air Soil Pollut.* **33**, 61–73 (1987).
- Schecher, W. D. & Driscoll, C. T. *Water Resour. Res.* **23**, 525–534 (1987).
- Hall, R. J. & Likens, G. E. *Nature* **292**, 329–331 (1981).
- Norton, S. A., Henriksen, A., Wathne, B. & Veidel, A. in *Proc. Acidification and Water Pathways*, 248–258 (Norwegian National Committee on Hydrology, Oslo, 1987).

- Henriksen, A., Wathne, B. M., Røgeberg, E. J. S., Norton, S. A. & Brakke, D. F. *Water Res.* (in the press).
- Norton, S. A. & Henriksen, A. *Vatten* **39**, 346–354 (1983).
- Misra, U. K., Blanchar, R. W. & Upchurch, W. J. *Soil Sci. Soc. Am. Proc.* **34**, 897–902 (1974).
- Bloom, P. R. & Grigal, D. F. *J. envir. Qual.* **12**, 489–495 (1985).
- Coleman, S. M. & Dethier, D. P. in *Rates of Chemical Weathering of Rocks and Minerals* (eds Coleman, S. M. & Dethier, D. P.) 1–20 (Academic, London, 1986).
- Tamm, C. O. & Hallbäck, L. *Water Air Soil Pollut.* **31**, 337–341 (1986).
- Falkengren-Grerup, U., Linnermark, N. & Tyler, G. *Chemosphere* **16**, 2239–2248 (1987).
- Hauhs, M. & Dise, N. *Geoderma* (in the press).
- Røgeberg, E. J. S. & Henriksen, A. *Vatten* **41**, 48–53 (1985).
- Cronan, C. S. & Aitken, G. R. *Geochim. cosmochim. Acta* **49**, 1697–1705 (1985).
- Brakke, D. F., Norton, S. A. & Henriksen, A. *Nature* **329**, 432–434 (1987).
- Henriksen, A., Brakke, D. F. & Norton, S. A. *Envir. Sci. Technol.* (in the press).
- Reuss, J. O. & Johnson, D. W. *Acid Deposition and the Acidification of Soils and Waters* (Springer, New York, 1986).
- Reuss, J. O., Cosby, B. J. & Wright, R. F. *Nature* **329**, 27–32 (1987).
- Wright, R. F. & Cosby, B. J. *Atmos. Envir.* **21**, 727–730 (1987).
- Dillon, P. J., Reid, R. A. & Girard, R. *Water Air Soil Pollut.* **31**, 59–65 (1986).
- Keller, W., Pitblado, J. R. & Conroy, N. I. *Water Air Soil Pollut.* **31**, 765–774 (1986).
- Battarbee, R. W. *et al.* *Nature* **332**, 530–532 (1988).
- Hauhs, M. & Wright, R. F. *Acid Deposition: Reversibility of Soil and Water Acidification—A Review* (Air Pollution Report No. 11, Commission of the European Communities, Brussels, 1988).
- Norwegian State Pollution Control Authority (SFT) *Monitoring of long-range transported polluted air and precipitation. Annual report 1986* Rep. no. 296/87 (SFT, Oslo, 1987) (in Norwegian).

# Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice

Christopher C. Goodnow, Jeffrey Crosbie, Stephen Adelstein, Thomas B. Lavoie\*, Sandra J. Smith-Gill\*, Robert A. Brink, Helen Pritchard-Briscoe, John S. Wotherspoon, Robert H. Loblay, Kathy Raphael†, Ronald J. Trent & Antony Basten

Clinical Immunology Research Centre, University of Sydney, New South Wales 2006, Australia

\* Laboratory of Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

† Division of Animal Production, CSIRO, Prospect, New South Wales, Australia

*Immunological tolerance has been demonstrated in double-transgenic mice expressing the genes for a neo-self antigen, hen egg lysozyme, and a high affinity anti-lysozyme antibody. The majority of anti-lysozyme B-cells did not undergo clonal deletion, but were no longer able to secrete anti-lysozyme antibody and displayed markedly reduced levels of surface IgM while continuing to express high levels of surface IgD. These findings indicate that self tolerance may result from mechanisms other than clonal deletion, and are consistent with the hypothesis that IgD may have a unique role in B-cell tolerance.*

AUTOIMMUNITY is a rare event despite the genetic potential of every individual to mount immune responses to autologous antigens. This is due to the fact that the immune system somehow 'learns' to recognize these antigens as self and to tolerate them<sup>1</sup>. The way in which self tolerance occurs is thought to depend on a range of factors including the time of exposure to antigen during ontogeny, the tissue site of expression, concentration, and the structure of the antigen<sup>2-5</sup>, although the relative importance of each of these factors in induction of tolerance to individual antigens remains obscure. It has been postulated that self-reactive lymphocytes, which might otherwise cause responses, either undergo clonal deletion from the pre-immune repertoire or remain in a functionally silent form as a result of early or repeated exposure to antigen (clonal anergy), contact with antigen-specific suppressor T cells or regulatory elements of an idiotype network<sup>2,6,7</sup>. These mechanisms are not mutually exclusive; for example, clonal deletion within the thymus or bone marrow could be reinforced by antigen-specific suppression in secondary lymphoid tissue.

The continuing controversy surrounding the mechanism of self tolerance stems largely from the difficulty in determining the fate of anti-self lymphocytes which, like lymphocytes with specificity for foreign antigens, are generated at very low frequencies in the pre-immune repertoire. This applies particularly to the high affinity cells believed to be important for pathogenic autoimmune responses. Thus although self-reactive T cells and B cells have been detected in secondary lymphoid tissue<sup>2,8,9</sup>, their functional significance remains unclear for two reasons. First, their detection has relied on the use of complicated functional assays or read-out systems that pick up the more frequent lymphocytes with very low affinity for antigen; secondly the cells represent an unknown fraction of the total output of anti-self lymphocytes from thymus or bone marrow.

One elegant approach to following the fate of self-reactive lymphocytes has involved the identification of naturally occurring situations in which the frequency of T cells specific for certain polymorphic cellular antigens is relatively high and can be measured directly by anti-V $\beta$  monoclonal antibodies<sup>10-12</sup>. In such situations the presence of self antigen clearly led to clonal deletion of 70-97% of anti-self T cells. However, this strategy is difficult to extend to other T cell specificities where there is less correlation between V-region usage and antigen specificity and where the frequency of anti-self T cells is lower; nor can it be readily applied to B cells where the V-region repertoire is

much larger<sup>13,14</sup>.

An alternative strategy is to generate high frequencies of anti-self lymphocytes in the pre-immune repertoire, through the introduction of rearranged T-cell receptor<sup>15,16,55</sup> or immunoglobulin genes<sup>17-19</sup> into the germline of transgenic mice. Here, we have followed this approach by generating two kinds of transgenic mice. The first kind carries the gene for a normally foreign protein, hen egg lysozyme (HEL), expressed in such a way that the transgenic antigen is recognized as self and leads to tolerance in anti-lysozyme T and B cells. The second kind of transgenic mouse carries rearranged immunoglobulin heavy and light chain genes encoding a high-affinity anti-lysozyme antibody, resulting in high frequencies of B-cells expressing this single receptor specificity. By mating the two kinds of transgenic mice, 'double-transgenic' offspring carrying lysozyme and immunoglobulin transgenes were produced. The anti-lysozyme B cells in these double-transgenic mice were functionally silenced, but not through clonal deletion. Rather they no longer responded efficiently even after adoptive transfer to mice not previously exposed to HEL and, intriguingly, were found to display greatly reduced levels of cell surface IgM but unchanged levels of IgD. This experimental strategy has the dual advantage of providing an *in vivo* system of tolerance induction wherein antigenic variables can be rigorously controlled and the fate of self-reactive high-affinity B cells can be followed with precision.

## Lysozyme-transgenic mice

Hen-egg lysozyme was selected as the transgenic self antigen for the following reasons: first, the biochemistry and antigenic structure of lysozyme are well defined; second, DNA clones for lysozyme and large amounts of protein were available; finally lysozyme is non-toxic *in vivo* and is therefore unlikely to exert nonspecific effects on the immunological milieu of recipient mice<sup>20,21</sup>. A genomic fragment containing the coding exons of the chicken lysozyme gene was linked to the mouse metallothionein I promoter<sup>22</sup> (Fig. 1a). The choice of the metallothionein promoter was based on the fact that it is transcriptionally active during embryonic, fetal and adult life, and can be induced to higher levels of transcription by heavy metals such as zinc at any of these stages<sup>23</sup>. The hybrid gene was microinjected into fertilised, inbred C57BL/6 eggs. Eight of 41 mice (20%) carried the transgene, and five of the eight (63%) had measurable amounts of lysozyme in their serum (0.8-28 ng ml<sup>-1</sup>, without induction by zinc).



## Tolerance in lysozyme-transgenic mice

To assay for tolerance to lysozyme in the metallothionein/lysozyme transgenic mice we selected one line, ML-5, with the highest basal levels of serum lysozyme (for ML-5 offspring: median = 17.3 ng ml<sup>-1</sup>, interquartile range 13–20 ng ml<sup>-1</sup>) and a single integrated copy of the transgene. It was not, however, possible to test the ML-5 mice for tolerance by direct challenge with lysozyme since C57BL/6 mice respond poorly to this antigen due to a recessive *Ir* gene effect<sup>24</sup>. C57BL/6 ML-5 mice which were hemizygous for the transgene were therefore mated with CBA (dominant high responder) mice, yielding litters of F<sub>1</sub> animals (responders) containing equal numbers of transgenic individuals and nontransgenic controls.

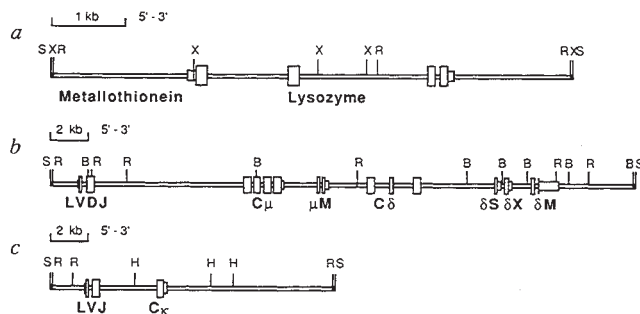
As shown in Fig. 2a, little or no anti-lysozyme antibody was detected by ELISA in the sera of transgenic mice from (CBA × B6)F<sub>1</sub> litters immunized with lysozyme in adjuvant (HEL/CFA). By contrast, high titres were obtained in all of the nontransgenic littermates. In theory the absence of detectable antibody in the transgenic mice could have been due to masking of the antibody by circulating lysozyme rather than tolerance, even though the low (nanomolar) serum lysozyme concentration in ML-5 mice is two orders of magnitude below that required to inhibit binding of anti-lysozyme antibodies in ELISA. To exclude this possibility formally, the number of cells actively secreting anti-lysozyme antibody was measured by plaque-forming cell (PFC) assay using lysozyme-coupled red blood cells (Fig. 2b). No direct (IgM) or indirect (IgG) plaque-forming cells were detected in any of the transgenic mice, whereas large numbers of IgG anti-lysozyme secreting cells were present in all of the nontransgenic littermates. Thus transgenic mice from the ML-5 line are indeed tolerant to lysozyme.

The lack of an anti-lysozyme antibody response in ML-5 mice could involve inactivation of lysozyme specific T cells, B cells, or both. Tolerance within the T-cell compartment was examined in an *in vitro* proliferative assay. Lysozyme-primed lymph node cells were obtained from (CBA × B6)F<sub>1</sub> offspring of the ML-5 transgenic line and the proliferative response was measured following restimulation with soluble lysozyme *in vitro*. Cells from the nontransgenic donors proliferated strongly after exposure to lysozyme as expected, whereas the response from cells of transgenic littermates was minimal (Fig. 2c). In addition the 'transgenic' T cells acted as a poor source of help in an adoptive anti-hapten antibody response (data not shown).

These results showed that tolerance in ML-5 mice affects lysozyme-specific T-cells, but did not indicate whether the B-cell compartment was unresponsive as well. To resolve this issue, C57BL/6 rather than F<sub>1</sub> offspring from the ML-5 line were immunized with lysozyme coupled to sheep red blood cells (HEL/SRBC), the latter acting as a foreign carrier to provide an alternative source of T-cell help. When the anti-lysozyme antibody response was measured by ELISA a 50-fold reduction in specific antibody titre was observed in the transgenic mice compared with their nontransgenic littermates (Fig. 2d). In contrast, there was no difference in the anti-SRBC titres between the same two groups of mice (data not shown). Comparisons of the number of anti-lysozyme plaque-forming cells in the spleens of these mice showed a similar but less dramatic trend, nevertheless confirming an effect of tolerance on the B-cell compartment (Fig. 2e). As tolerance is thought to affect high-affinity cells preferentially<sup>2,3</sup>, the most plausible explanation for the discrepancy between the two assays was that the plaque assay detected more low affinity antibodies than the ELISA. This possibility was confirmed by demonstrating a requirement for higher concentrations of soluble lysozyme to inhibit PFC lysis from transgenic donors than from their nontransgenic littermates (Fig. 2f).

## Immunoglobulin-transgenic mice

The findings in ML-5 mice confirmed the feasibility of inducing tolerance to transgene-encoded antigens<sup>25–27</sup>. In addition they



**Fig. 1** Structure of microinjected lysozyme and immunoglobulin gene constructs. *a*, Hybrid metallothionein/lysozyme gene. *b*, Anti-lysozyme immunoglobulin heavy chain gene. *c*, Anti-lysozyme kappa light chain gene. The thin open bars indicate flanking or intron sequences, 5' and 3' untranslated regions (5'UT, 3'UT) are represented by open boxes, and the coding regions by wider boxes. Polylinker sequences at either end are shown in black. Relevant restriction sites are indicated: B, *Bam*HI; H, *Hind*III; R, *Eco*RI; S, *Sal*I; X, *Xba*I.

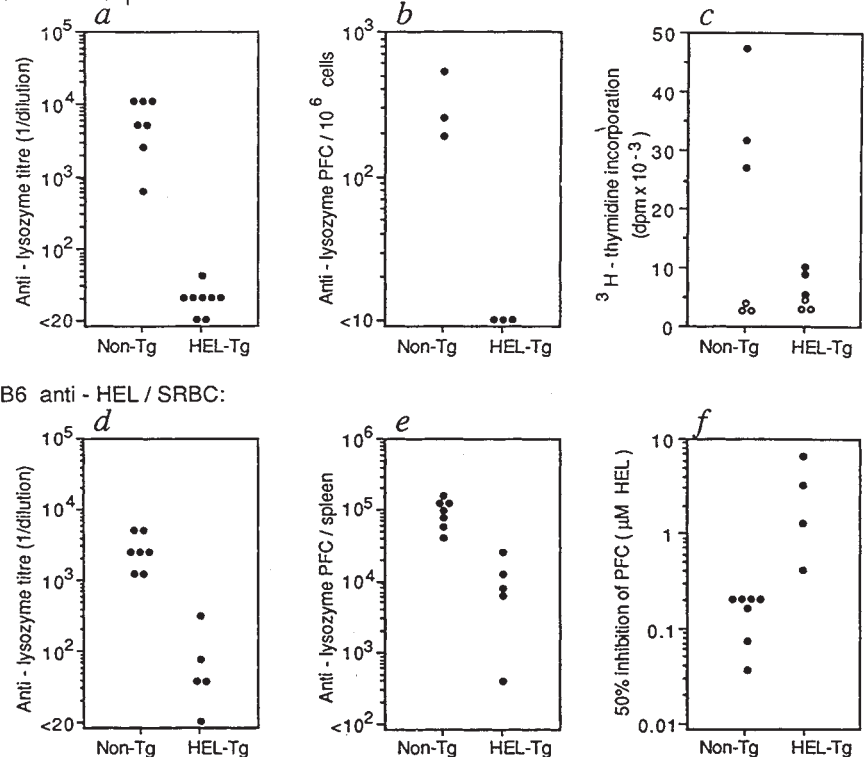
**Methods.** Gene constructs were prepared using standard techniques<sup>41</sup>. *a*, The metallothionein (MT) promoter was derived from pMK (ref. 22) as a 1.9 kilobase (kb) *Eco*RI–*Bgl*II fragment. The lysozyme gene was derived from pLysG (ref. 42). An *Xba*I site was generated from the *Bgl*II site at nucleotide +68 in the metallothionein 5' UT, and from the *Rsa*I site at nucleotide +15 in the lysozyme 5' UT, by subcloning into the polylinkers of pGC2 (ref. 43) and pUC19 respectively. The two genes were then assembled into the *Eco*RI site of the plasmid pSVG-gpt, a derivative of pSV2ΔH-gpt (ref. 44), in which the *Bam*HI site had been destroyed by Klenow fill-in and the 300 base pair (bp) *Eco*RI–*Pst*I fragment replaced by a polylinker. The polylinker sites, in order from the end nearest the *Amp*<sup>R</sup> gene, are *Cl*aI, *Xho*I, *Sal*I, *Xba*I, *Bgl*II, *Nco*I, *Eco*RI, *Sac*I, *Kpn*I, *Sma*I, *Bam*HI, *Xba*I, *Sal*I, *Pst*I. *b*, The LVDJH segments were derived from a genomic clone of the HyHEL10 heavy chain gene, λEMBL4-402-1 (T. B. Lavoie, *et al.*, manuscript in preparation) as a 3.9 kb *Eco*RI partial digest fragment, which was joined to a 12.8 kb *Eco*RI fragment from λgt-WES-IgH-701 (ref. 45) containing an undelimited  $\mu$ -switch region,  $\mu$ -constant region (C $\mu$ ) and  $\mu$ -membrane exons ( $\mu$ M). The  $\delta$ -constant region (C $\delta$ ) and  $\delta$ -secretory exon ( $\delta$ S) were derived from a 7.1 kb *Eco*RI–*Bam*HI partial digest fragment subcloned by Drs V. T. Oi and J. Dangel from  $\gamma$ CH28-257.3 (ref. 46). The *Sal*I site in the C $\delta$ - $\delta$ S intron was inactivated by Klenow fill-in and religation. The  $\delta$ X and  $\delta$ -membrane ( $\delta$ M) exons were isolated on a 7.1 kb *Bam*HI partial digest fragment of λCH30-372.6 (ref. 47). The three segments were assembled together to form a 31.3 kb V<sub>H</sub>–C $\mu$ –C $\delta$  gene between the *Eco*RI and *Bam*HI sites of pSVG-gpt. *c*, A 4.7 kb *Eco*RI partial–*Hind*III fragment containing L $\kappa$  and V $\kappa$  exons was derived from a genomic clone of the HyHEL10 kappa light chain gene, λEMBL3-10.15K2-A (T. B. Lavoie *et al.*, in preparation). The  $\kappa$ -constant region and 3' flanking DNA were derived from a genomic clone, λCh4-ABPC4-11A, of the productive kappa chain gene in the plasmacytoma ABPC4 (ref. 48), as an 11.7 kb *Hind*III partial–*Eco*RI fragment. The two parts were assembled into the *Eco*RI site of pSVG-gpt. All three hybrid genes (*a*, *b* and *c*) were excised with *Sal*I and purified by agarose gel electrophoresis followed by caesium chloride density gradient centrifugation and dialysis<sup>41,49</sup>. The metallothionein/lysozyme gene was injected at a concentration of 2  $\mu$ g ml<sup>-1</sup>; the heavy and light chain genes were mixed in an equimolar ratio and co-injected at a total DNA concentration of 2  $\mu$ g ml<sup>-1</sup>. Fertilized eggs were obtained from matings of C57BL/6J mice maintained at the University of Sydney. Microinjection, embryo transfer, and screening of pups by Southern blot of tail DNA was performed as described<sup>49</sup>.

provided further support for the concept that tolerance can influence the B-cell as well as the T-cell compartment and that this effect is most noticeable in the higher affinity B cells<sup>3</sup>. Owing to their low frequency, however, it was not possible to determine the fate of these cells nor to distinguish between the possible mechanisms of tolerance operating, such as clonal deletion, clonal anergy or suppression. To overcome this problem we

**Fig. 2** Anti-lysozyme B- and T-cell responses in lysozyme-transgenic mice and non-transgenic littermates. Litters of (CBA×B6)<sub>F</sub><sub>1</sub> or C57BL/6 offspring from hemizygous ML-5 transgenic mice were immunized with hen egg lysozyme in adjuvant (HEL/CFA) or lysozyme coupled to sheep red blood cells (HEL/SRBC) respectively. Responses were measured in non-transgenic (non-Tg) and transgenic (HEL-Tg) littermates. *a* and *d*, Serum anti-lysozyme antibody titres. *b* and *e*, Number of indirect (IgG) plaque forming cells (PFC) against lysozyme-coupled red cells per 10<sup>6</sup> lymph node cells (*b*) or per spleen (*e*). *c*, *In vitro* proliferation of lysozyme-primed lymph node cells cultured with (closed circles) or without (open circles) lysozyme in the culture medium. *f*, Concentration of soluble lysozyme producing 50% inhibition of IgG anti-HEL/horse red blood cell (HEL/HRBC) plaques.

**Methods.** Two to four month old (CBA×B6)<sub>F</sub><sub>1</sub> offspring were injected intraperitoneally (i.p.) with 100 µg purified hen egg lysozyme in complete Freund's adjuvant (HEL/CFA) and sera were collected 24 days later. Two to four month old C57BL/6 offspring were given two i.v. injections of 10<sup>8</sup> lysozyme-coupled sheep red blood cells<sup>50</sup>, 49 days apart; sera were collected 4 days after the second injection. Antibodies were detected by ELISA on lysozyme-coated flat-bottomed 96-well microtitre trays as described<sup>28</sup>, except that goat anti-mouse Ig/alkaline phosphatase (Sigma) was used as the second step, the reaction developed with *p*-nitrophenyl phosphate, and absorbance (*A*<sub>405</sub>) at 405 nm determined on a Titertek Multiscan. Titres were determined as the last twofold serial dilution with an *A*<sub>405</sub> > 2 s.d. above that of pre-immune sera. Anti-sheep red blood cell antibody titres were also determined in sera from HEL/SRBC immunized mice by haemagglutination and were not significantly different between the transgenic and nontransgenic groups (data not shown). For plaque-forming cell responses in (CBA×B6)<sub>F</sub><sub>1</sub> animals, 10 µg lysozyme in CFA was injected subcutaneously at the base of the tail and the inguinal lymph nodes collected 10 days later. PFC's were measured in the C57BL/6 offspring at the same time as the serum titres were determined, 4 days after boosting with HEL/SRBC. Indirect plaques were developed<sup>50,51</sup> against HEL/SRBC from HEL/CFA-immunized mice, or against HEL/HRBC from mice immunized with HEL/SRBC. Specificity was confirmed in parallel using unconjugated red cells, and few or no plaques were detected in any experiment. The concentration of soluble lysozyme required to inhibit 50% of the PFC in individual animals was determined by adding 10-fold dilutions of HEL to the plaquing mixture, plotting the number of plaques against [HEL], and interpolating the 50% inhibition value relative to the number of plaques in the absence of HEL. *In vitro* proliferative responses from inguinal lymph nodes of mice primed 10 days previously at the base of the tail with 10 µg HEL/CFA were performed as described<sup>52</sup>, in round-bottomed 96-well microtitre plates (Linbro) containing 10<sup>5</sup> lymph node cells and 4×10<sup>5</sup> irradiated (2,000 rad) syngeneic spleen cells per well, with or without 100 µg ml<sup>-1</sup> lysozyme added to the culture medium. Lymph node cells from each animal were cultured separately in triplicate for 72 h before pulsing for 18 h with <sup>3</sup>H-thymidine; the mean <sup>3</sup>H-thymidine incorporation per well is shown.

(B6×CBA)<sub>F</sub><sub>1</sub> anti-HEL:



increased the frequency of high affinity anti-lysozyme B cells by introducing rearranged immunoglobulin genes into a parallel set of transgenic mice. The hybridoma HyHEL10 (ref. 28) was selected as the source of lysozyme-specific heavy and light chain genes because the antibody secreted by this hybridoma has a high affinity for lysozyme ( $K_a = 2 \times 10^9 \text{ M}^{-1}$ , M. E. Denton and H. A. Sheraga, personal communication) and the interaction of the antibody with its epitope has been characterized in detail (ref. 28 and E. Padlan, E. Silvertown, S. Sheriff, G. Cohen, S. Smith-Gill and D. Davies, unpublished results).

Genomic clones encoding the HyHEL10 heavy and light chains (T. B. Lavoie *et al.*, manuscript in preparation) were used for the immunoglobulin transgene constructs. The  $\gamma 1$ -constant region of the HyHEL10 heavy chain gene was replaced by a segment containing the  $\mu$  and  $\delta$  constant region genes (Fig. 1b). The rationale for using the complete  $\mu$ - $\delta$  locus was to ensure that the progression from expression of IgM alone to co-expression of IgM plus IgD which is observed during normal B-cell maturation<sup>29</sup>, could occur in B cells expressing the transgene. The  $\mu$  and  $\delta$  genes were derived from BALB/c (IgH<sup>a</sup> allotype) mice. This allowed transgene-derived heavy chains to be distinguished by anti-allotypic monoclonal antibodies from products of the endogenous C57BL/6 (IgH<sup>b</sup> allotype) heavy chain genes.

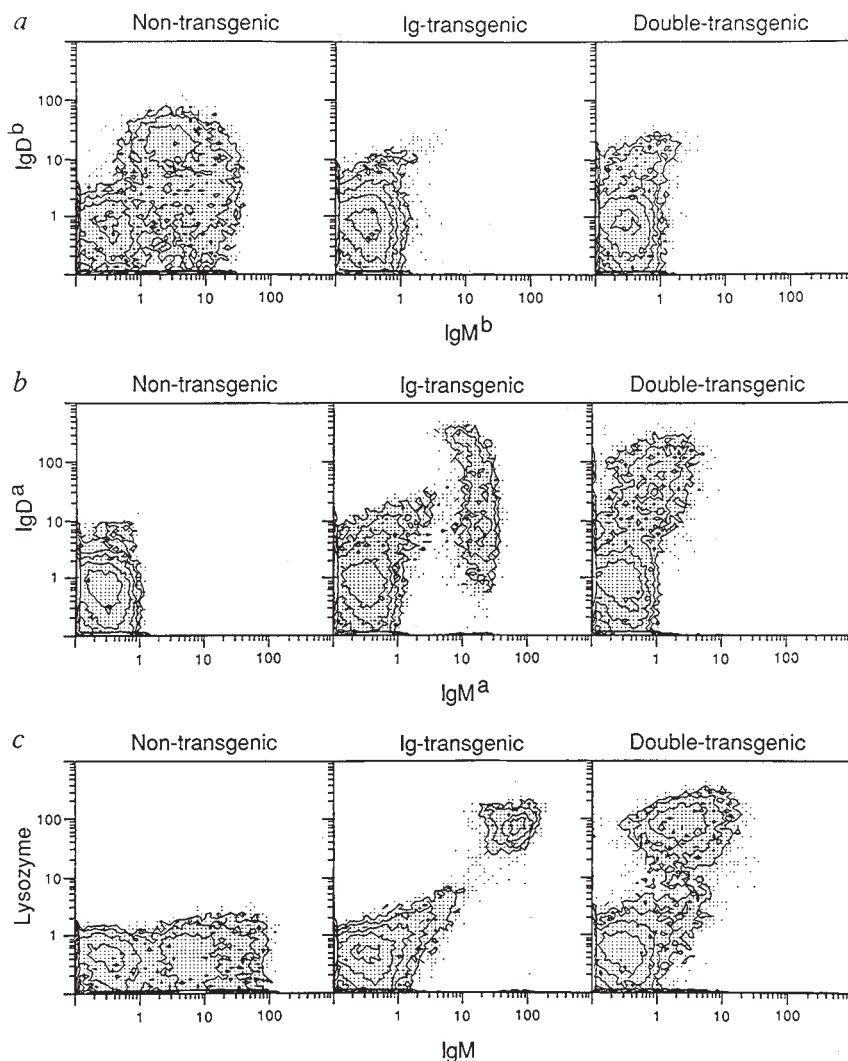
The light and heavy chain genes were co-injected into C57BL/6 eggs, as co-injection usually results in the integration of both genes into a single chromosomal locus<sup>19</sup> and co-segregation during breeding. Seven of 47 mice born were transgenic (15%); six of these carried copies of both light and heavy chain transgenes whereas the seventh contained only the heavy chain. High serum titres of anti-lysozyme antibodies were detected in all six animals with both immunoglobulin genes. The heavy chain-only founder had very low levels of anti-lysozyme activity but high levels of a-allotype IgM (data not shown). These data confirm the importance of both transgene-derived chains for high-affinity lysozyme binding.

One line, MD-3, was bred from a male immunoglobulin (Ig)-transgenic founder which carried a single integration site containing multiple copies of heavy and light chain transgenes. To determine the frequency of anti-lysozyme B-cells in MD-3 offspring, spleen cells were stained with monoclonal anti-allotype antibodies or a sandwich of soluble lysozyme followed by a noncompeting monoclonal antibody to lysozyme. Two-colour FACS analysis showed that over 90% of the splenic B-cells expressed transgene-derived a-allotype heavy chains as IgM and/or IgD (Fig. 3b), and that 60–90% of these cells bound lysozyme (Fig. 3c and Table 1). Furthermore, only 1–6% of spleen cells bearing the endogenous b-allotype IgM or IgD were



**Fig. 3** Two-colour FACS analyses of spleen cells from non-transgenic, Ig-transgenic and double-transgenic mice. Spleen cells were stained with monoclonal antibodies specific for IgM and IgD of the b-allotype to identify cells carrying endogenous C57BL/6 heavy chains (a), or monoclonal antibodies specific for IgM and IgD of the a-allotype to identify cells bearing transgene-encoded IgM or IgD (b). To identify lysozyme-binding B cells (c), cells were stained with an excess of unlabelled lysozyme followed by a complementary biotinylated anti-lysozyme monoclonal antibody, and counterstained with a rat monoclonal antibody to all mouse IgM. Contour maps show correlated fluorescence in arbitrary relative fluorescence units, with fluorescein (FITC) emission on the x-axis and phycoerythrin (PE) emission on the y-axis and contour intervals at 5, 10, 20, 40 and 80 counts per square of a 4×4 channel grid.

**Methods.** Spleen cell suspensions from 6-week-old transgenic (MD-3.2, MD-3.6) or nontransgenic mice were stained<sup>31,32</sup> with the following monoclonal antibodies: anti-IgM<sup>b</sup>, AFS-78.25/FITC; anti-IgD<sup>b</sup>, AF6-122.2/biotin; anti-IgM<sup>a</sup>, DS-1/FITC; anti-IgD<sup>a</sup>, AMS-15.1—biotin; anti-IgM, 331.12/FITC. For lysozyme binding, cells were stained with 200 ng ml<sup>-1</sup> HEL followed by 200 ng ml<sup>-1</sup> HyHEL9/biotin<sup>28</sup>. Binding of biotinylated reagents was revealed with streptavidin/PE (Becton-Dickinson). Four-parameter data were collected for 3×10<sup>4</sup> cells on a FACS 440 with linear amplification of forward and side-scatter and logarithmic amplification of fluorescence emission<sup>34</sup>. Each profile represents fluorescence data from list mode files after gating out dead cells, granulocytes and macrophages on the basis of forward and side-scatter. The percentage positively staining cells was determined from cutoffs set at  $x = 1.15$ ,  $y = 10.0$  (a and b), or at  $x = 1.15$ ,  $y = 3.8$  (c).



**Table 1** B-cell frequency and phenotype in nontransgenic, Ig-transgenic and double-transgenic mice

| Mouse       | Sex | Age<br>(weeks) | Transgene |    | % IgM <sup>b+</sup><br>and/or IgD <sup>b+</sup> | % IgM <sup>a+</sup><br>and/or IgD <sup>a+</sup> | HEL <sup>+</sup> cells |                         | Relative IgM fluorescence |                        |
|-------------|-----|----------------|-----------|----|-------------------------------------------------|-------------------------------------------------|------------------------|-------------------------|---------------------------|------------------------|
|             |     |                | HEL       | Ig |                                                 |                                                 | %                      | no. (×10 <sup>6</sup> ) | HEL <sup>+</sup> cells    | HEL <sup>-</sup> cells |
| Spleen      |     |                |           |    |                                                 |                                                 |                        |                         |                           |                        |
| Non-Tg      | M   | 6              | —         | —  | 55.5                                            | 0.2                                             | 0.1                    | 0.08                    | —                         | 0.32, 5.5              |
| MD-3.2      | M   | 6              | —         | +  | 3.7                                             | 29.2                                            | 16.1                   | 9.3                     | 62                        | 0.35                   |
| MD-3.5      | M   | 6              | —         | +  | 4.4                                             | 27.0                                            | 17.8                   | 11.7                    | 56                        | 0.33                   |
| MD-3.6      | M   | 6              | +         | +  | 5.8                                             | 26.3                                            | 34.2                   | 13.3                    | 2.3                       | 0.31                   |
| MD-3.9      | F   | 6              | +         | +  | 4.7                                             | 19.4                                            | 25.0                   | 10.8                    | 2.0                       | 0.32                   |
| Non-Tg      | M   | 8              | —         | —  | 48.3                                            | 0.7                                             | 0.2                    | 0.42                    | —                         | 0.28, 3.5              |
| MD-3.10     | M   | 8              | —         | +  | 1.2                                             | 8.3                                             | 7.0                    | 13.3                    | 39                        | 0.30                   |
| MD-3.11     | M   | 8              | +         | +  | 1.0                                             | 3.9                                             | 2.8                    | 6.7                     | 1.6                       | 0.28                   |
| Lymph node  |     |                |           |    |                                                 |                                                 |                        |                         |                           |                        |
| Non-Tg      | M   | 8              | —         | —  | 13.2                                            | 0.3                                             | 0.0                    | —                       | —                         | 0.31                   |
| MD-3.10     | M   | 8              | —         | +  | 1.3                                             | 19.4                                            | 21.5                   | —                       | 22                        | 0.40                   |
| MD-3.11     | M   | 8              | +         | +  | 0.7                                             | 6.4                                             | 6.9                    | —                       | 1.1                       | 0.32                   |
| Bone marrow |     |                |           |    |                                                 |                                                 |                        |                         |                           |                        |
| Non-Tg      | M   | 8              | —         | —  | 9.2                                             | 0.7                                             | 0.2                    | —                       | —                         | 0.30                   |
| MD-3.10     | M   | 8              | —         | +  | ND                                              | 3.6                                             | 4.6                    | —                       | 22                        | 0.31                   |
| MD-3.11     | M   | 8              | +         | +  | ND                                              | 2.9                                             | 1.8                    | —                       | 1.6                       | 0.28                   |

Transgenic littermates were genotyped as described in Fig. 4. Spleen, mesenteric lymph node, and bone marrow cell suspensions were stained and analysed by two-colour FACS as described in Fig. 3. The total number of lysozyme-binding cells (HEL<sup>+</sup>) = % positive × number of cells per spleen. Relative IgM fluorescence was determined in arbitrary relative fluorescence units from the peak (modal) green fluorescence (331.12/FITC) within the lysozyme-binding (HEL<sup>+</sup>) and non-binding (HEL<sup>-</sup>) populations. Where two peaks were present, as in the HEL<sup>-</sup> cells from nontransgenic spleen, the values for both peaks are shown. FACS analysis of spleen cells from mice carrying only the lysozyme-transgene gave identical results to nontransgenic C57BL/6 mice. ND, not determined.

detected (Fig. 3a and Table 1) and these showed very weak staining. The low frequency of B-cells bearing endogenous heavy chains most likely reflects allelic exclusion of endogenous gene rearrangement by the already rearranged transgenes<sup>30</sup>, however the degree of inhibition in MD-3 mice is far greater than has been previously observed.

Compared with the marked abnormalities reported previously in B cells from  $\mu$ -transgenic mice<sup>31</sup>, the B cells in MD-3 mice proved to be relatively normal by several criteria. Thus there was only a moderate reduction in their frequency and numbers in spleen and lymph node, although the reduction may become more marked with age (Table 1). Second, they lacked Ly-1 and Mac-1 but did express B220, J11D, and Ia in amounts identical to the predominant population of splenic B cells from the nontransgenic controls (data not shown). Finally their Ig-phenotype, IgM<sup>hi</sup>, IgD<sup>lo-hi</sup> (Fig. 3b), although different from the IgM<sup>lo</sup>, IgD<sup>hi</sup> phenotype characteristic of splenic B cells from nontransgenic adult mice, does have a normal counterpart in that it is very similar to that of B-cells in the spleen of 1–2 week-old mice<sup>32</sup>.

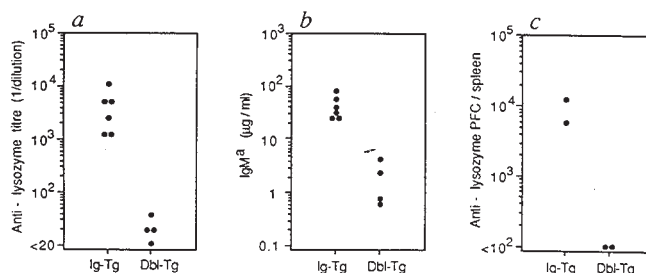
### Tolerance in double-transgenic mice

Double-transgenic mice bearing the immunoglobulin transgenes and the lysozyme transgene were produced by mating the MD-3 founder male with ML-5 females which were hemizygous for the lysozyme transgene. In the first two litters, 4/17 mice carried both antibody and lysozyme transgenes (double-transgenic), 6/17 carried only the immunoglobulin transgenes (Ig-transgenic), 6/17 carried only the lysozyme transgene (HEL-transgenic) and 1/17 was nontransgenic.

Spontaneous secretion of high serum titres of anti-lysozyme antibody, at levels comparable to those in the founder animals, was observed in the Ig-transgenic mice (Fig. 4a). By contrast, in the double-transgenic littermates, lysozyme-binding antibody was either undetectable or present in extremely low concentrations. When the serum concentration of all IgM antibodies bearing the transgenic a-allotype was measured in the double-transgenic mice (Fig. 4b) a similar although less marked reduction in titres was evident. Furthermore, the spleens of Ig-transgenic mice contained large numbers of anti-lysozyme plaque-forming cells (Fig. 4c) whereas none were detected in the spleens of double-transgenic mice, once again arguing against simple masking of the antibody by circulating lysozyme.

Two-colour FACS analysis was used to determine whether the lack of anti-lysozyme antibody secretion in double-transgenic mice was due to clonal deletion of B-cells expressing the transgenic antibody. Lymphocytes expressing transgene-encoded receptors were identified either by binding of monoclonal anti-a-allotype antibodies specific for transgenic IgM or IgD heavy chains (IgH<sup>a</sup>), or by binding of lysozyme itself. Representative fluorescence profiles are shown in Fig. 3 and the complete data are summarized in Table 1.

Analysis of double transgenic mice clearly demonstrated that the majority of B cells expressing the transgene-encoded receptor were not clonally deleted. At six weeks of age there was no reduction in the frequency or number of spleen cells expressing a-allotype heavy chains in the double transgenics, nor was there a decrease in the frequency or numbers of lysozyme-binding cells, compared with the Ig-transgenic littermates (Table 1). The amount of IgM expressed on the B cells was, however, dramatically and selectively reduced (Fig. 3b, c). The modal fluorescence of double-transgenic B cells stained with a rat monoclonal antibody that binds equally well to a- and b-allotype IgM was an average of 3.8% that of the B-cells in the control Ig-transgenic littermates and 42% that of splenic B-cells in nontransgenic mice (Table 1). In contrast, the levels of IgD were unaltered. The selectivity of the effect for IgM was further demonstrated by the lack of any detectable change in the expression of other surface molecules such as B220, J11D, La, Ly-1 or Mac-1. At eight weeks of age these differences in immunoglobulin



**Fig. 4** Anti-lysozyme antibody secretion in Ig-transgenic (Ig-Tg) and double-transgenic (Dbl-Tg) mice. *a*, Serum titres of anti-lysozyme antibody. *b*, Serum concentration of a-allotype IgM. *c*, No. of direct (IgM) anti-lysozyme plaque-forming cells per spleen.

**Methods.** Mice were genotyped for the lysozyme transgene by DNA dot-blot<sup>19</sup> and for immunoglobulin transgenes by Southern blot, and analysed at 4–6 weeks of age. Serum titres of anti-lysozyme antibody were determined as described in Fig. 2. IgM of the a-allotype was measured by inhibition ELISA using the RS-3.1 monoclonal antibody as described<sup>53</sup>, except that plates were coated with TEPC-183 (IgM<sup>a</sup>,  $\kappa$ ) myeloma protein, and RS-3.1 culture supernatant was used diluted 1/500 followed by sheep anti-mouse IgG-alkaline phosphatase (Southern Biotechnology Associates).

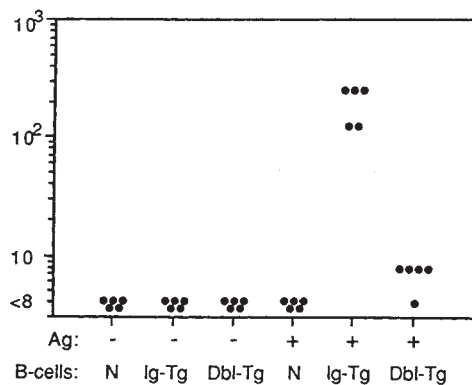
Direct (IgM) plaque-forming cells were assayed as for Fig. 2.

phenotype were still evident, and were observed in bone marrow and lymph nodes as well as in the spleen. A decline in the number and frequency of lysozyme-specific B cells in the spleens of double-transgenic versus Ig-transgenic littermates was, however, also apparent. This may be age-related and could reflect a lack of exogenous antigenic stimulation of lysozyme-specific B cells, an altered distribution amongst the lymphoid tissues, or terminal differentiation. But we cannot exclude the alternative possibility that a proportion of B cells may have been deleted (for example, those with high densities of surface IgM).

To exclude the possibility that the observed changes in secretion or surface expression of immunoglobulin in the double-transgenic mice might be due to production of significant amounts of lysozyme by the B cells, chimaeras were constructed by injecting MD-3 Ig-transgenic bone marrow cells into lethally irradiated ML-5 lysozyme-transgenic recipients or non-transgenic littermates. Analysis of the chimaeras three weeks later showed that the Ig-transgene expressing B-cells in the lysozyme-transgenic recipients displayed the same phenotypic changes and absence of spontaneous immunoglobulin secretion as seen in the double-transgenic mice.

In view of the absence of anti-lysozyme antibody secretion and the persistence of lysozyme-binding B-cells in an altered phenotypic form, the next step was to determine whether the B cells in the double-transgenic mice had undergone a persistent functional change. A small number ( $10^5$ ) of spleen cells from either a nontransgenic, Ig-transgenic, or double-transgenic mouse were injected into irradiated C57BL/6 recipients, together with an excess ( $5 \times 10^6$ ) of spleen cells from horse red blood cell-primed C57BL/6 mice as a source of helper T cells. Half of the recipients were injected with lysozyme covalently coupled to horse red blood cells at the time of transfer while the remainder served as controls. As shown in Fig. 5, moderate amounts of anti-lysozyme antibody were secreted in recipients of Ig-transgenic spleen cells, provided they were boosted with antigen (HEL-HRBC). In contrast, the titres in recipients of 'double-transgenic' spleen cells were 25-fold lower, and only marginally above the background levels in recipients of non-transgenic C57BL/6 spleen cells. The B cells from these double-transgenic mice therefore appear to respond poorly to stimulation with antigen in the presence of T-cell help even though the





**Fig. 5** Anti-lysozyme antibody secretion in adoptive transfer recipients. Groups of 5 sublethally irradiated C57BL/6 mice were reconstituted with  $10^5$  spleen cells from either non-transgenic (Non-Tg), Ig-transgenic (Ig-Tg), or double-transgenic (Dbl-Tg) mice, and  $5 \times 10^6$  spleen cells from horse red blood cell-primed C57BL/6 mice. The recipients were either not boosted (-) or boosted (+) with lysozyme-coupled horse red blood cells (Ag), and serum titres determined 7 days later.

**Methods.** Donors were unimmunised 8-week-old C57BL/6 non-transgenic or transgenic (MD-3.10, MD-3.11) males, or male C57BL/6 mice primed with 0.1 ml 20% HRBC i.p. 11 weeks previously. Recipients were 3–4 month old C57BL/6 males irradiated with 750 rads. Antigen boosted recipients received 0.1 ml 20% HEL/HRBC i.p.. Titres were determined as described for Fig. 2.

corresponding cells from Ig-transgenic mice expressing a transgene-encoded receptor can do so. Since the number of lysozyme-specific suppressor T cells present in  $10^5$  spleen cells is unlikely to be sufficient to produce suppression in the transfer recipients, these findings strongly suggest that the double-transgenic B cells have indeed undergone a functional change during development, particularly when taken in conjunction with the alterations in IgM/IgD phenotype.

## Discussion

The continuing controversy surrounding self tolerance has been due in large part to the lack of an ideal experimental model for analysing cellular events in tolerant hosts at the clonal and molecular levels. The use of double transgenic mice expressing the genes for a neo-self antigen (lysozyme) and a corresponding high affinity antibody (HyHEL10), offers a new approach to the problem and provides a reproducible way of controlling the many variables associated with the phenomenon of tolerance (see introduction). In the current paper particular attention is focussed on validating the model (Figs 1 and 2) and applying it to the study of the fate of self-reactive B cells and the role of IgM versus IgD in B cell regulation.

The finding that profound tolerance to lysozyme was maintained in the double transgenic mice without a major decrease in the frequency of anti-lysozyme B cells (Figs 3–5) stands in contrast to the evidence for clonal deletion of self-reactive T-cells with specificity for Mls<sup>a</sup> or I-E antigens<sup>10–12</sup>. At first sight these findings are suggestive of a possible dichotomy in the mechanisms of self tolerance for B cells and T cells, however other interpretations are possible. For example, the persistence of B cells in our tolerant mice also differs from previous data supporting deletion of immature B cells after exposure *in vitro* to multivalent ligands such as anti-immunoglobulin or complex antigens<sup>2,33,34</sup>. This discrepancy could in fact be explained if higher concentrations of lysozyme (induced for example by feeding zinc) or a multivalent form of lysozyme were shown to result in deletion, the reason being that lower concentrations of anti-immunoglobulin have been shown to induce B-cell anergy without deletion and the *in vitro* deletion

phenomenon seems to be dependent on multivalency<sup>2</sup>. Given the flexibility of the model with respect to antigen presentation, and the feasibility of manipulating the T-cell repertoire in a similar fashion<sup>15,16,55</sup> it should be possible to resolve whether the apparent differences in our data compared to that of others reflect a true dichotomy in tolerance between T cells and B cells or simply different responses to 'weak' versus 'strong' self-antigens common to both subsets.

A second possible interpretation of the changes observed in the double-transgenic mice is that anti-lysozyme B cells expressing high levels of surface IgM are deleted, and this allows the development of a variant population with low levels of IgM but unchanged levels of IgD. Experiments to determine the molecular basis and reversibility of the decreased surface IgM expression should help to resolve this issue.

The failure of anti-lysozyme B cells from the double transgenic mice to respond efficiently to antigen and T-cell help (Fig. 5) indicates that they may have undergone a persistent functional change. Further support for this conclusion comes from the striking alteration in surface immunoglobulin phenotype observed on FACS analysis (Table 1 and *vide infra*). On the other hand the findings do not allow us to distinguish between the various mechanisms which might bring about such a change. Thus a persistent functional change in the B cell could in theory be induced either by encounter with lysozyme at an early stage of development (clonal anergy), or by interaction with antigen specific suppressor T-cells, or regulatory elements of an idiotype network. Adoptive transfer of different subsets of cells at various stages of development together with the use of different lines of Ig-transgenic mice should help to distinguish between these alternatives.

The loss of surface IgM with persistent expression of IgD on the B cells of double-transgenic mice was a completely unexpected finding, and raises interesting questions as to the role of these immunoglobulin isotypes in tolerance induction. A negative correlation between surface IgD and susceptibility to tolerance has often been hypothesized because of the delayed appearance of IgD during B-cell maturation, and the sensitivity of immature or mature B-cells lacking IgD to inactivation by multivalent ligands *in vitro*<sup>35–37</sup>. On the other hand, the picture has been complicated by the absence of significant differences in the signalling properties of IgM and IgD for generation of intracellular second messengers, antigen uptake and processing, or B-cell activation *in vitro*<sup>38</sup>. The striking change in IgM/IgD receptor ratio induced in the double-transgenic mice (Fig. 3) provides further evidence for a differential role of IgD and IgM in B-cell tolerance, although the persistent expression of IgD on the tolerant B cells appears to contradict the hypothesis that IgD confers resistance to tolerance<sup>35</sup>. The precise role of the two isotypes, in particular whether the change in receptor status is a cause or an effect of the functional silencing in the double-transgenic B cells, will clearly be an important direction for future studies.

At first sight it is surprising that the selective downregulation of IgM demonstrated here has not been observed in studies of tolerance *in vitro*. Anti-IgM antibodies or antigen have been shown either to cause irreversible modulation of all surface immunoglobulin (clonal abortion) or to affect function but not surface immunoglobulin expression (clonal anergy)<sup>2,33,34</sup>. Selective downregulation of IgM on anergic B cells may have been overlooked, however, since IgD is not expected to be present on very immature B cells and has not been analysed. Intriguingly, the closest counterpart to the phenotypic changes in the B cells from the double-transgenic mice is found in normal B cell maturation. Indeed, the IgM<sup>hi</sup>/IgD<sup>lo-hi</sup> phenotype of the B cells in the immunoglobulin-only transgenic littermates is similar to the phenotype of immature B cells which predominate in the spleens of one to two week-old mice, while the IgM<sup>lo</sup>, IgD<sup>hi</sup> phenotype of the double-transgenic B cells is similar to the mature B cells predominant in spleens of older mice<sup>32</sup> (Fig. 3).

The experimental strategy of mating antigen-transgenic mice with antigen-receptor transgenic mice offers many advantages. Most importantly, it allows tolerance to be studied at the level of the whole animal, and yet provides a bridge between *in vivo* processes and the growing body of molecular and *in vitro* immunological data. Second, the induction of tolerance by a transgenic self antigen more accurately reproduces the physiological situation than does administration of exogenously synthesized antigens, particularly in view of the evidence that exogenous antigens may not be presented in association with class I major histocompatibility molecules<sup>39,40</sup>. Third, experimental manipulation of the structure of the transgenic self antigen, tissue site of synthesis, timing of expression, or concentration can be achieved either by altering the microinjected construct or simply by comparing mice with different integration sites<sup>26,27</sup>. Finally, the transgenic antigen receptor can be varied, for example by altering the fine specificity, isotype, or transmem-

brane domains encoded by the microinjected immunoglobulin genes.

We thank Dr Paul Lalor for advice and the gift of anti-B-cell and anti-allotype monoclonal antibodies, Dr Donna Sieckmann for the DS-1 antibody, Dr Eberhardt Weiler for antibody RS-3.1, and Tracy Anderson and Pat Gregory for technical assistance. We also thank Drs Andrew Wilkes and Richard Palmiter for plasmid pMK, Dr Günther Schütz for plasmids plysG and p1023, Dr Tasuku Honjo for  $\lambda$ gt-WES-IgH-701, Dr Vernon Oi for plasmid pSV2- $\delta$ , Dr Philip Tucker for  $\lambda$ Ch30-372.6, and Drs Jerry Adams and Liz Webb for  $\lambda$ Ch4-ABPC4-11A. We particularly thank Dr Mark Davis and Phil Patten for encouragement and advice. C.C.G. and S.A. are supported by NHMRC research scholarships, and R.A.B. is supported by a CPRA. The project was supported by grants from the National Health and Medical Research Council, and the Commonwealth Tertiary Education Commission Centres of Excellence Committee.

Received 2 May; accepted 8 July 1988.

1. Burnet, F. M. *The Clonal Selection Theory of Acquired Immunity* (Vanderbilt University Press, Nashville, 1959).
2. Nossal, G. J. V. *A. Rev. Immun.* **1**, 33-62 (1983).
3. Siskind, G. W. in *Fundamental Immunology* (ed. Paul, W. E.) 537-558 (Raven, New York, 1984).
4. Dresser, D. W. & Mitchison, N. A. *Adv. Immun.* **8**, 129-181 (1968).
5. Weigle, W. O. *Adv. Immun.* **16**, 61-122 (1973).
6. Basten, A., Loblay, R. H., Trent, R. J. & Gatenby, P. A. in *Recent Advances in Clinical Immunology* (ed. Thompson, R. A.) 33-63 (Churchill Livingstone, New York, 1980).
7. Jerne, N. K. *Ann. Immun. (Inst. Pasteur)* **125C**, 373-389 (1974).
8. Ternynck, T. & Avrameus, S. *Immun. Rev.* **94**, 99-112 (1986).
9. Wood, P. J., Socarras, S. & Streilein, J. W. *J. Immun.* **139**, 3236-3244 (1987).
10. Kappler, J. W., Roehm, N. & Marrack, P. *Cell* **49**, 273-280 (1987).
11. Kappler, J. W., Staerz, U. W., White, J. & Marrack, P. *Nature* **332**, 35-40 (1988).
12. Macdonald, H. R. *et al. Nature* **332**, 40-45 (1988).
13. Davis, M. M. & Bjorkman, P. J. *Nature* (in the press).
14. Tonegawa, S. *Nature* **302**, 575-581 (1983).
15. Ivars, F. *et al.* in *The T-cell Receptor* (eds Kappler, J. W. & Davis, M. M.) 187-197 (Liss, New York, 1988).
16. Uematsu, Y. *et al. Cell* **52**, 831-841 (1988).
17. Grosschedl, R., Weaver, D., Baltimore, D. & Costantini, F. *Cell* **38**, 647-658 (1984).
18. Rusconi, S. & Köhler, G. *Nature* **314**, 330-334 (1985).
19. Storb, U. *et al. J. exp. Med.* **164**, 627-641 (1986).
20. Osserman, E. F., Canfield, R. E. & Beychok, S. (eds) *Lysozyme* (Academic, New York, 1974).
21. Benjamin, D. C. *et al. A. Rev. Immun.* **2**, 67-101 (1984).
22. Palmiter, R. D., Chen, H. Y. & Brinster, R. L. *Cell* **29**, 701-710 (1982).
23. Palmiter, R. D. & Brinster, R. L. *A. Rev. Genet.* **20**, 465-499 (1986).
24. Gammon, G. *et al. Immun. Rev.* **98**, 53-73 (1987).
25. Chisari, F. V. *et al. Science* **230**, 1157-1160 (1985).
26. Adams, T. E., Alpert, S. & Hanahan, D. *Nature* **325**, 223-228 (1987).
27. Faas, S. J., Pan, S., Pinkert, C. A., Brinster, R. L. & Knowles, B. B. *J. exp. Med.* **165**, 417-427 (1987).

28. Smith-Gill, S. J., Lavoie, T. B. & Mainhart, C. R. *J. Immun.* **133**, 384-393 (1984).
29. Blattner, F. R. & Tucker, P. W. *Nature* **307**, 417-422 (1984).
30. Weaver, D., Costantini, F., Imanishi-Kari, T. & Baltimore, D. *Cell* **42**, 117-127 (1985).
31. Herzenberg, L. A. *et al. Nature* **329**, 71-73 (1987).
32. Hardy, R. R., Hayakawa, K., Parks, D. R. & Herzenberg, L. A. *Nature* **306**, 270-272 (1983).
33. Sidman, C. L. & Unanue, E. R. *Nature* **257**, 149-151 (1975).
34. Raff, M. D. *et al. J. exp. Med.* **142**, 1052-1064 (1975).
35. Vitetta, E. S. & Uhr, J. W. *Immun. Rev.* **37**, 50-88 (1977).
36. Kettman, J. R., Cambier, J. C., Uhr, J. W., Ligler, F. & Vitetta, E. S. *Immun. Rev.* **43**, 69-95 (1979).
37. Scott, D. W., Venkataraman, M. & Jandinski, J. J. *Immun. Rev.* **43**, 241-280 (1979).
38. Möller, G. (ed.) *Immun. Rev.* **95** (1987).
39. Townsend, A. R. M., Bastin, J., Gould, K. & Brownlee, G. G. *Nature* **324**, 575-577 (1986).
40. Morrison, L. A., Lukacher, A. E., Braciale, V. L., Fan, D. P. & Braciale, T. J. *J. exp. Med.* **163**, 903-921 (1986).
41. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1983).
42. Matthias, P. D., Renkawitz, R., Grez, M. & Schütz, G. *EMBO J.* **1**, 1207-1212 (1982).
43. Myers, R. M., Lerman, L. S. & Maniatis, T. *Science* **229**, 242-247 (1985).
44. Mulligan, R. C. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2072-2076 (1981).
45. Nikaido, T., Nakai, S. & Honjo, T. *Nature* **292**, 845-848 (1981).
46. Liu, C.-P., Tucker, P. W., Mushinski, J. F. & Blattner, F. R. *Science* **209**, 1348-1353 (1980).
47. Cheng, H.-L., Blattner, F. R., Fitzmaurice, L., Mushinski, J. F. & Tucker, P. W. *Nature* **296**, 410-415 (1982).
48. Webb, E., Adams, J. M. & Cory, S. *Nature* **312**, 777-779 (1984).
49. Hogan, B., Costantini, F. & Lacy, E. *Manipulating the Mouse Embryo* (Cold Spring Harbor Laboratory, New York, 1986).
50. Kipp, D. E. & Miller, A. in *Selected Methods in Cellular Immunology* (eds Mishell, B. B. & Shiigi, S.) 103-105 (Freeman, San Francisco, 1980).
51. Cunningham, A. J. & Szentgyörgyi, A. *Immunology* **14**, 599 (1968).
52. Corradin, G., Etlinger, H. M. & Chiller, J. M. *J. Immun.* **119**, 1048-1053 (1977).
53. Schüppel, R., Wilke, J. & Weiler, E. *Eur. J. Immun.* **17**, 739-741 (1987).
54. Adams, E., Wotherspoon, J. S., Hellqvist, L. & Basten, A. *Immun. Cell Biol.* **65**, 25-33 (1987).
55. Kisielow, P., Blüthmann, H., Staerz, U. D., Steinmetz, M. & von Boehmer, H. *Nature* **333**, 742-748 (1988).

## LETTERS TO NATURE

### Self-excited cosmic string dynamos

David N. Spergel\*, William H. Press†  
& Robert J. Scherrer†

\* Institute for Advanced Study, Princeton, New Jersey 08540, USA

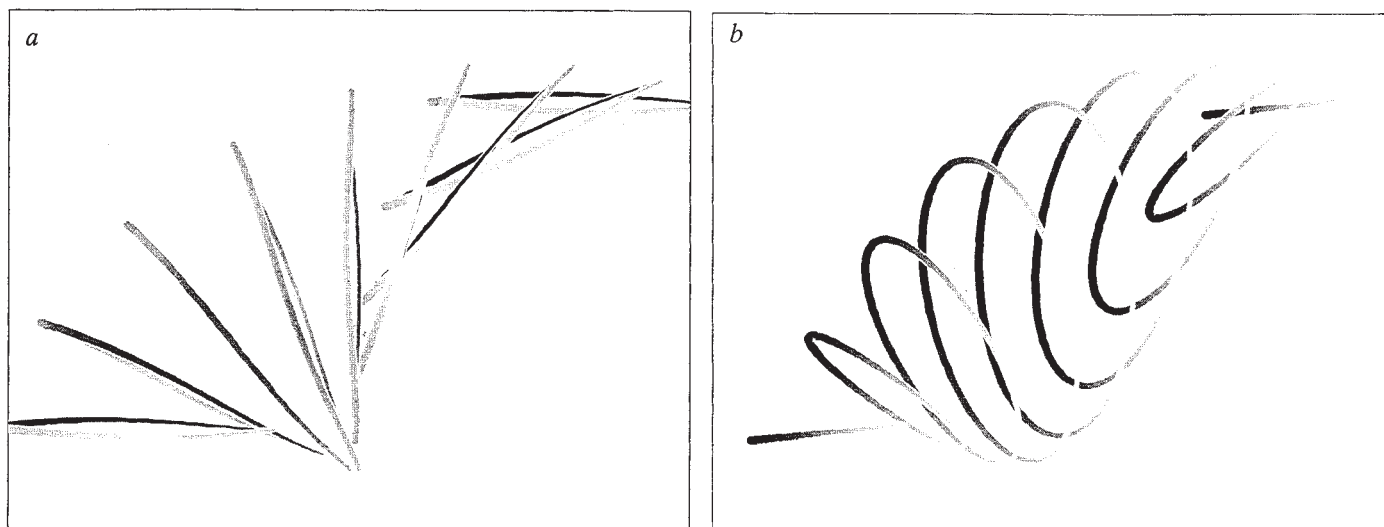
† Department of Physics and Harvard-Smithsonian Center for Astrophysics, Harvard University, Cambridge, Massachusetts 02138, USA

Cosmic strings, topological remnants of the earlier universe, arise in many grand unified theories. Witten<sup>1</sup> showed that under some conditions these strings can behave like superconductors: current is carried by massless charge carriers, which can be either fermions or bosons, that move at the speed of light along the string. These strings are superconductors in the sense that any induced constant current (d.c.) will persist. The persistence is guaranteed by topo-

logical index theorems<sup>1</sup>, but no such theorem exists for alternating currents (a.c.), which can be damped by radiation or (as we report here) can even grow exponentially through dynamo self-interaction. This mechanism converts the mechanical energy of the string into electromagnetic fields.

When an oscillating cosmic string loop crosses a stationary magnetic field line, two sets of charge carriers are produced. One moves to the left and the other, carrying the opposite charge, moves to the right. When the string loop contracts and crosses the field line in the opposite direction, two sets of charge carriers are again produced, but the signs of the carriers are reversed. This oscillation produces an a.c. current, but no d.c. current. Because these charges move along the string loop with the same period as the oscillation of the loop itself, they will return to the same point when the string next crosses the field line. It has previously been shown<sup>2,3</sup> that there are growing a.c. modes in a cosmic string that is oscillating in an external inhomogeneous magnetic field. These modes grow only linearly in time, but their existence suggested the possibility of modes growing exponentially due to self-interaction<sup>3</sup>.





**Fig. 1** The mechanical evolution of a string trajectory that (a) admits growing modes and (b) does not. Each loop in the picture represents the string at a different phase in its oscillation. The darker, thicker lines indicate depth in dimension perpendicular to the page.

We have investigated the behaviour of self-interacting test currents on isolated, oscillating loops of string<sup>4</sup>. 'Test' means that the current is small enough not to act back on the string's mechanical motion. We determine the electromagnetic fields produced by charges moving along the string and calculate the currents produced on the string by these fields. We have explored the behaviour of currents in the family of string oscillation modes described by Chen *et al.*<sup>5-7</sup> and in the  $m=1$ ,  $n=2$  trajectories described by Burden<sup>8</sup>. These trajectories are solutions to the wave equation. We find that on the highly symmetric Burden strings, all a.c. modes are damped, but in the less symmetric trajectories of Chen *et al.*<sup>7</sup> a large fraction of the solutions contain at least one growing a.c. mode. This directly demonstrates the existence of superconducting cosmic string dynamos. (Details of the calculation will be given elsewhere<sup>4</sup>).

The rotating loop-like string trajectory shown in Fig. 1a admits growing modes; the one illustrated in Fig. 1b does not. In the loop of Fig. 1a the charge carriers rearrange themselves so that there is always a net leftward current at one tip and a net rightward current at the other. There is always a net positive charge on the near side of the string and a net negative charge on the far side.

If a weak current is generated on a superconducting string with a growing a.c. mode, it will grow exponentially: the current generates a field that acts back on the string and produces additional current. The characteristic growth time for these modes is  $1/\alpha$  ( $\sim 137$ ) times the string oscillation period, much shorter than the timescale for the dynamic evolution of strings. Any external seed field could produce this initial current. The current will grow either until it becomes large enough to affect the dynamics of string motion or until it is suppressed by other effects, such as collisions between charge carriers along the string, vacuum polarization effects or interaction with the surrounding plasma.

While bosonic strings can carry a substantial current without ill effect, some fermionic strings may have their superconductivity suppressed. In certain theories in which strings have fermionic charge carriers, the a.c. current may be dissipated through weak resistive effects<sup>9</sup>. Another potential obstacle to field generation arises if the fermionic charge carriers have finite mass along the string<sup>10</sup>.

When the current exceeds an electric charge per electron Compton wavelength, it can dissipate its energy through pair creation. The alternating current will generate in places an  $E$ -like field strong enough to polarize the vacuum and create electron-positron pairs, which will dissipate some of the energy in the electromagnetic field.

In the early universe, the string is not oscillating in a vacuum, but in a dense plasma, an excellent conductor which is opaque to low-frequency electromagnetic radiation. The interaction of this plasma with the string will affect the growth of the magnetic field.

While both plasma effects and vacuum breakdown may fight the growth of currents along the string, our calculations of strings *in vacuo* show that there is no symmetry or conservation law that forbids the exponentially increasing conversion of a string's mechanical energy into electromagnetic modes. In complex systems, if equilibration of energy between modes is not forbidden, it will often occur. The existence of a mechanism for transferring the kinetic energy of the string into electromagnetic energy suggests that seed currents can grow, even in complicated situations. These current-carrying strings can have profound effects on their surrounding environment: the formation of large scale structure<sup>11</sup>, the powering of quasars<sup>12</sup>, the emission of gamma-ray bursts<sup>13</sup>, the creation of ultra-high-energy cosmic rays<sup>14</sup>, and the radiation of synchrotron emission<sup>15</sup>.

W.H.P. and R.J.S. acknowledge support from NSF. D.N.S. is supported by NSF and by a New Jersey High Technology Grant 88-240090-2.

Received 27 April; accepted 1 July 1988.

1. Witten, E. *Nucl. Phys.* B249, 557-592 (1985).
2. Spergel, D. N., Piran, T. & Goodman, J. J. *Nucl. Phys.* B291, 847-875 (1987).
3. Aryal, M., Vilenkin, A. & Vachaspati, T. *Phys. Lett.* B194, 25-29 (1987).
4. Spergel, D. N., Press, W. H. & Scherrer, R. J. *Phys. Rev. D.* (submitted).
5. Kibble, T. W. B. & Turok, N. *Phys. Lett.* B116, 141-143 (1982).
6. Turok, N. *Nucl. Phys.* B242, 520-541 (1984).
7. Chen, A. L., DiCarlo, D. A. & Hotes, S. A. *Phys. Rev.* D37, 863-868 (1988).

8. Burden, C. J. *Phys. Lett.* B164, 277-281 (1985).
9. Barr, S. N. & Matheson, A. M. *Phys. Rev.* D36, 2905-2914 (1987).
10. Hill, C. T. & Widrow, L. M. *Phys. Lett.* B189, 17-22 (1987).
11. Ostriker, J. P., Thompson, C. & Witten, E. *Phys. Lett.* B180, 231-239 (1986).
12. Vilenkin, A. & Field, G. B. *Nature* 326, 772-773 (1986).
13. Babul, A., Paczynski, B., & Spergel, D. N. *Astrophys. J.* 316, L49-L54 (1987).
14. Hill, C. T., Schramm, D. N. & Walker, T. P. *Phys. Rev.* D36, 1007-1016 (1987).
15. Chudnofsky, E. M., Field, G. B., Spergel, D. N. & Vilenkin, A. *Phys. Rev.* D34, 944-950 (1986).

# Optical observations of the eclipsing binary radio pulsar PSR1957+20

J. van Paradijs\*, J. Allington-Smith†, P. Callanan‡, P. A. Charles§, B. J. M. Hassall‡, G. Machin‡, K. O. Mason||, T. Naylor¶ & A. P. Smale||

\*Astronomical Institute 'Anton Pannekoek', University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands

†Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

‡Department of Astrophysics, Oxford University, South Parks Road, Oxford OX1 3RQ, UK

§Royal Greenwich Observatory, Apartado 321,

3870 Santa Cruz de La Palma, Tenerife, Canary Islands

||Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Dorking, Surrey RH5 6NT, UK

¶IUE Observatory, European Space Agency, Apartado 54065, 28080 Madrid, Spain

The recently discovered<sup>1</sup> 1.6 ms binary radio pulsar PSR1957+20 shows radio eclipses whose duration indicates that the occulting body is substantially larger than the Roche lobe of the low-mass ( $\sim 10^{-2} M_{\odot}$ ) companion star. This suggests<sup>1-4</sup> that this companion is evaporating through the action of a strong pulsar energy flux<sup>5</sup>. An optical counterpart has been detected which shows brightness variations in phase with the 9.2 h orbital cycle<sup>6</sup>. We have obtained optical charge coupled device (CCD) images which show that the counterpart is one component of a close visual pair separated by  $\sim 0.7$  arcsec. At maximum both are equally bright with combined *V* magnitude of 19.9, while at minimum PSR1957+20 is invisible. From spectroscopic observations we find that the contaminating star is a normal G star. The spectrum of PSR1957+20 shows intermittent H $\alpha$  emission. We confirm that the optical brightness of PSR1957+20 varies in phase with the radio Doppler velocity curve<sup>1</sup>, and find that the amplitude is probably more than 3 magnitudes, minimum light coinciding with the radio eclipse. The optical light curve is consistent with heating of a hydrogen-rich low-mass white dwarf by high-energy radiation from the nearby millisecond pulsar.

The observations were made with the 4.2 m William Herschel Telescope of the Observatorio del Roque de los Muchachos between 5 and 26 June 1988. CCD images in the *B*, *V* and *R* bands were obtained during the nights of 5, 21, 23 and 24 June UT, using the RGO CCD camera equipped with a GEC3 dye-coated chip with enhanced blue response. The camera was mounted for direct imaging at one of the Nasmyth *f*/11 foci

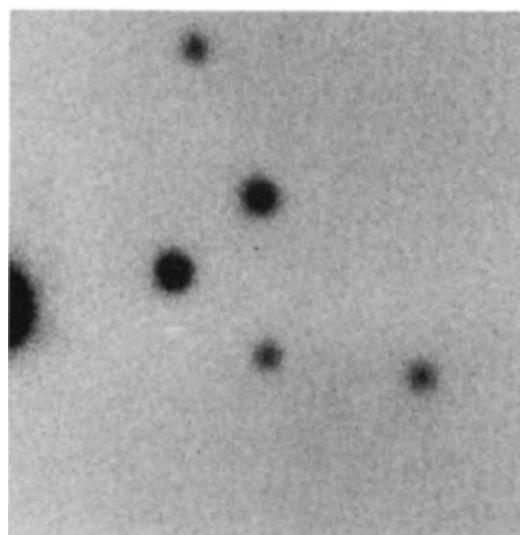
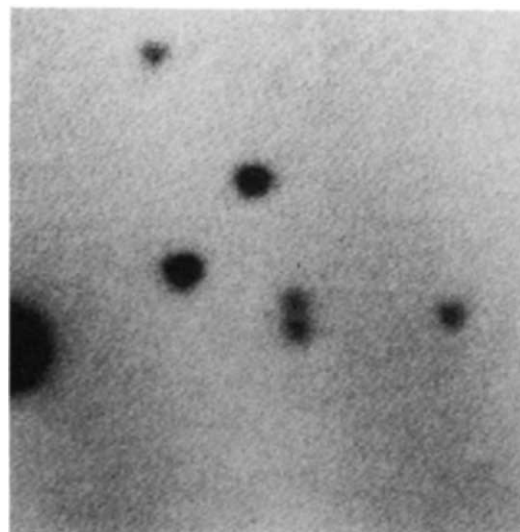


Fig. 1 Details of two CCD frames taken on 23 June 1988, UT 01:51 (top) and UT 04:34 (bottom), at orbital phases 0.92 and 0.22, respectively (using the ephemeris from ref. 1). The weakening of the optical counterpart of PSR1957+20 near the radio eclipse (phase 0.25) is clearly visible. The field shown measures about  $16 \times 16$  arcsec. The directions toward the North and East are approximately toward the lower-left and lower-right corners of the pictures, respectively. The star to the left of centre was also put on the slit during spectroscopic observations.

Table 1 Image and spectral data

| CCD observations |             |                   |                      |
|------------------|-------------|-------------------|----------------------|
| UT date (1988)   | UT (start)  | Exposure time (s) | No. of frames        |
| 5 June           | 01:43-03:46 | 300, 400          | 17 (B)               |
|                  | 04:40-05:30 | 300, 500          | 8 (R)                |
| 21 June          | 03:51-04:03 | 300               | 3 (V)                |
|                  | 04:10       | 300               | 1 (B)                |
| 23 June          | 00:05-04:56 | 200, 300          | 27 (V), 5 (B), 6 (R) |
| 24 June          | 00:07-04:24 | 200               | 27 (V), 2 (B), 1 (R) |

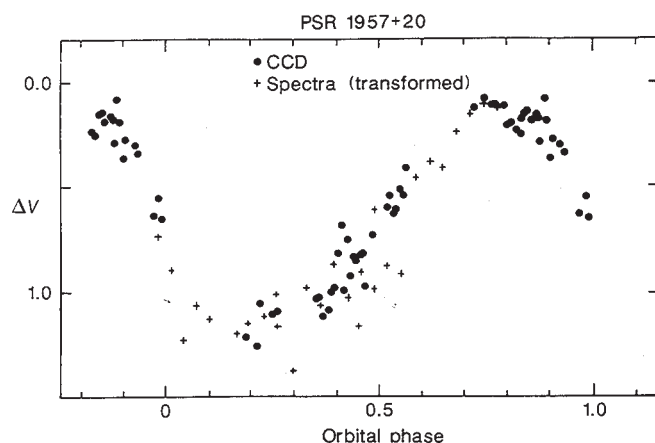
  

| FOS observations |             |                |                |
|------------------|-------------|----------------|----------------|
| UT (1988)        | UT (start)  | Phase interval | No. of spectra |
| 8 June           | 03:31-04:36 | 0.88-0.96      | 4              |
| 15 June          | 01:45-04:17 | 0.98-0.26      | 10             |
| 17 June          | 02:10-04:14 | 0.26-0.49      | 8              |
| 19 June          | 01:44-04:42 | 0.45-0.78      | 11             |

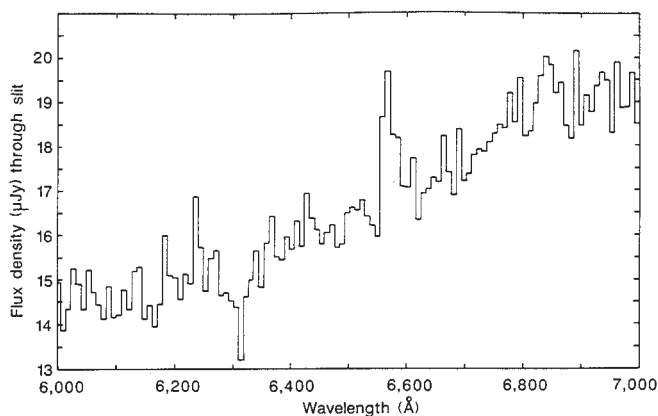
giving a plate scale of 0.1 arcsec per  $22 \mu$  pixel and a field of view of  $58 \times 38$  arcsec. 97 CCD frames were obtained (see Table 1). Images obtained under good seeing conditions show that the image of the optical counterpart is double (see Fig. 1), with the two components separated by about 0.7 arcsec (see below).

We performed aperture photometry of the combined image of PSR1957+20 (star A) and the contaminating star (star B), using the Pandora software package (A. Pickles, private communication). The brightness of a star is determined from the sum signal in a fixed circular aperture centred on the stellar image (the aperture was set at 4 arcsec diameter). This signal is corrected for sky background, determined in a concentric ring around the source aperture, from which cosmic rays and stars





**Fig. 2** Light curve (relative  $V$  magnitude) for the combined image of the optical counterpart of PSR1957+20 and the nearby contaminating star. From observations of standard stars we find that  $\Delta V = 0$  corresponds to  $V = 19.75$ . Dots and crosses indicate, respectively, results obtained from the CCD images and from the spectra.



**Fig. 3** Average of 8 spectra (1,000 s integrations each) of PSR 1957+20, taken on 17 June 1988 between UT 02:10 and 04:34, showing weak  $H\alpha$  emission.

are filtered out. From a comparison of the signal from PSR1957+20 (combined image of the pair) with the results for stars on the same frame of brightness similar to the pulsar system we estimate that the r.m.s. accuracy of a single measurement of the brightness of the binary pulsar is  $\sim 0.04$ – $0.09$  magnitude. A calibration of the CCD photometry was obtained from observations of standard stars, taken from the Kitt Peak list.

During four nights (covering all phases of the orbit) we performed spectroscopy on PSR1957+20 using the highly efficient RGO-Durham Faint Object Spectrograph<sup>7</sup> (FOS) at the Cassegrain focus of the WHT. This is a fixed-format spectrograph which provides coverage of the spectrum between 3,600 and 5,000 Å (second order, at 5 Å per pixel) and between 5,000 and 9,300 Å (first order, at 9 Å per pixel). The detector is again a dye-coated GEC CCD which provides a spatial resolution of 1.2 arcsec per pixel along the 20 arcsec slit, whose width was set at 1.0 arcsec. The integration time of each spectrum was 1,000 s. 33 spectra were obtained (see Table 1). The FOS data were analysed in near-real time using an interactive FOS software package which performed the extraction of the (curved) spectrum from the CCD frame, sky signal subtraction, correction for atmospheric extinction and atmospheric emission bands, and a flux calibration, using observations of the standard star HZ 44 (ref. 8). Except for the first night the spectra were taken with the slit oriented to transmit the light of both PSR1957+20 (both components of the optical pair) and a nearby star of similar brightness. This allowed us to derive a crude differential magnitude estimate of PSR1957+20 by integrating the fluxes of the two objects over the wavelength range 5,000 to 6,000 Å. These relative magnitudes deviate systematically from the results of the CCD photometry. This is probably due to different settings on the slit of the pair and the comparison star, and a difference in the passbands. We have corrected the spectroscopic data by scaling the amplitude of the magnitude variation (by a factor of 1.37) and applying a zero point shift (of 0.43 magnitude). The resulting transformed data have been included in the  $V$  band light curve of the pair.

The light curve (see Fig. 2) has a single broad minimum near the phase of radio eclipse, with a maximum about half a cycle later. The amplitude of the  $V$  curve is  $\sim 1.0$  magnitude and its average is  $V \approx 20.3$ . The light curve is possibly asymmetric, with a decrease to minimum that is somewhat steeper than the rise to maximum, but because the decreasing part of the light curve is determined mainly by spectroscopic data we do not consider that this asymmetry is firmly established.

To estimate the contribution of star B we analysed some CCD frames, obtained under good seeing conditions, using the Pan-

dora software (with an assumed 1 arcsec diameter aperture) and DAOPHOT<sup>9</sup>. We find a separation between the two stars of  $\sim 0.7$  arcsec; at phase 0.92 (ephemeris from ref. 1) PSR1957+20 contributes a fraction  $0.50 \pm 0.06$  to the  $V$  brightness of the pair. This implies that near minimum light star A is at least a factor 10 fainter in  $V$  than star B. The amplitude of the modulation of the  $V$  magnitude of PSR1957+20 (after correcting for star B) is  $> 3$  magnitudes. At maximum the contribution of PSR 1957+20 corresponds to a magnitude  $V = 20.4$ .

From a comparison of the  $B$  and  $V$  data obtained around maximum and minimum light we find that the colour index ( $B - V$ ) of the pair does not change significantly with brightness. From these data, and the fact that at minimum star A does not contribute significantly, we infer for PSR1957+20 alone a value ( $B - V$ ) =  $0.8 \pm 0.1$  (at maximum light). The lack of variation in  $B - V$  is confirmed by a comparison of the  $B$  data obtained on 5 June (which cover the phase interval 0.79–0.99) with the corresponding part of the  $V$  light curve.

The average spectrum near the phase of radio eclipse is dominated by star B. We have compared this spectrum with spectra of bright stars of known spectral type, taken with the FOS during twilight. From the lack of strong Balmer lines and of other strong features we estimate that the spectrum of the contaminating star is of type G. The average spectrum obtained during the night of 17 June showed  $H\alpha$  in emission (see Fig. 3) with an equivalent width of 4 Å ( $\pm 10\%$ ). After correction for the signal from star B the equivalent width of  $H\alpha$  is about 12 Å.

The large amplitude of the optical light curves of the PSR 1957+20 system, its phasing with respect to the radio eclipses and its approximately sinusoidal shape indicate that most of the optical light from the system is due to a heating effect of the secondary, most probably by irradiation with high-energy photons and/or particles from the millisecond radio pulsar. To estimate the optical luminosity of the system and the size of the optical emitter, we need to know its bolometric flux, interstellar reddening and distance.

We have estimated the distance,  $d$ , to PSR1957+20 from its dispersion measure<sup>1</sup>, DM, using the average of the ratio  $d/DM$  ( $33.7 \pm 3.8$  pc per unit DM, one standard deviation error) for pulsars at similar galactic longitudes<sup>10</sup>. This gives  $d \approx 900$  pc, with an uncertainty of about a factor of two on both sides<sup>11</sup>.

The relation between distance and visual interstellar extinction  $A_V$  in the Cygnus region varies strongly with galactic longitude<sup>12</sup>. For directions near that of PSR1957+20 ( $l^{II} = 60^\circ$ ,  $b^{II} = -3^\circ$ )  $A_V$  rises on average by  $\sim 1$  magnitude per kpc up to a distance of 2 kpc. However, interstellar extinction is patchy

and substantial deviations from this average relation occur<sup>12</sup>. If we take this into account we find that available data weakly constrain the extinction of PSR1957+20 to be in the range 0.3 to 3.6 magnitude. An upper limit on  $A_V$  may be inferred from the integrated column density of neutral hydrogen,  $N(\text{H I})$ , as determined from 21 cm radio data<sup>13</sup>, which for the 0.5° beam containing PSR1957+20 is  $3.7 \times 10^{21} \text{ cm}^{-2}$ . Since CO observations<sup>14</sup> indicate that molecular hydrogen does not contribute much to the column density in this direction, we have used the relation<sup>15</sup> between  $A_V$  and total column density  $N(\text{H I} + \text{H}_2)$  to infer an upper limit to  $A_V$  of 2.5 magnitude. The corresponding range in interstellar reddening  $E_{B-V}$  is 0.1 to 0.85 magnitude. With these ranges in distance and interstellar reddening the intrinsic colour index  $(B-V)_0$  and the absolute visual magnitude  $M_V$  range between 0.0 and 0.7, and between 6.2 and 11.5, respectively.

The above constraints on the interstellar reddening lead to a lower limit of the distance of the contaminating G type star of 2.5 kpc making it unlikely that this star is related to the binary pulsar.

Our knowledge of the reprocessing of the high-energy emission of a radio pulsar in a low-mass companion is inadequate to allow a trustworthy estimate of the resulting spectral energy distribution (in particular in the ultraviolet), which is required to determine the bolometric flux. While realizing the limitations we have decided to use stellar energy distributions to make this conversion. Using Popper's<sup>16</sup> relations between  $(B-V)$ ,  $M_V$ , and radius for main sequence stars we find that the bolometric luminosity of the PSR1957+20 system ranges between  $\sim 0.02$  and  $\sim 0.4 L_\odot$ ; the corresponding radius of the emitting region is between  $\sim 0.05$  and  $\sim 0.15 R_\odot$  (here we have used reasonable combinations of small/large distance and low/high interstellar reddening). As an alternative we have assumed that the energy distribution of the PSR1957+20 system is that of a low-mass X-ray binary (which is also dominated by reprocessing of high-energy radiation, in this case X rays). This assumption is only meaningful if the interstellar reddening is on the high side of the interval derived above. We then find<sup>17</sup> that the radius of the emitting region is  $\sim 10$  times smaller than that of a typical accretion disk in a low-mass X-ray binary, that is,  $\sim 0.1 R_\odot$ .

These estimates of the luminosity and size of the optically emitting region in PSR1957+20 have a large uncertainty, but they suffice to show that this region is fairly small,  $\sim 0.1 R_\odot$ . This is consistent with the idea that the optical emission originates from reprocessing of high-energy radiation from the millisecond pulsar in a low-mass ( $\sim 0.01 M_\odot$ ) white dwarf with a hydrogen-rich composition<sup>18</sup> (as suggested by the presence of H $\alpha$  emission in some of our spectra of the PSR1957+20 system).

We have estimated the fraction of the pulsar flux impinging on the companion that is transformed into optical light, using the above assumption that the spectral energy distribution is that of a normal star. We take a distance of  $2.5 R_\odot$  between the binary components and assume that the luminosity of PSR 1957+20 is similar to that of the single 1.55-ms pulsar PSR 1937+21 (ref. 19):  $L \sim 10^{36} \text{ erg s}^{-1}$ . Then the flux impinging on the companion is equivalent to an effective temperature of  $\sim 14,500 \text{ K}$ . The range in  $(B-V)_0$  derived above corresponds to effective temperatures between  $\sim 5,600 \text{ K}$  and  $\sim 9,500 \text{ K}$ , implying fractions of pulsar luminosity converted into optical % ultraviolet emission of between  $\sim 2$  and  $\sim 15\%$ . This result is consistent with the idea that a substantial fraction of the impinging pulsar flux does not lead to heating of the companion (see, for example, ref. 5).

Finally we can constrain the inclination of the orbital plane of PSR1957+20 from the lower limit on the amplitude of the optical brightness variations (a factor of 10). We assumed that (1) all optical light is the result of reprocessing of pulsar emission in the companion star; (2) the heated parts of the companion star radiate uniformly and isotropically; (3) the non-irradiated

parts of the companion do not emit optical light (see, however, ref. 20). As extreme cases we have taken a spherical companion star irradiated by a parallel beam, and one which fills its Roche lobe. For the latter case we have estimated the variation of the fraction of the heated surface visible to the observer, using a numerical model for calculating optical light curves of X-ray binaries<sup>21</sup> and assuming a mass ratio 0.015. For both cases we find that the inclination angle is larger than  $55^\circ$ .

The William Herschel Telescope is operated at La Palma by the Royal Greenwich Observatory in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias. We thank Drs J. H. Taylor and A. S. Fruchter for providing accurate coordinates of PSR1957+20, Dr H. Bloemen for information on interstellar reddening, and Mr M. Heemskerk for computations of heating light curves. T.N. is affiliated to the Astrophysics Division, Space Sciences Division, Space Sciences Department.

Received 8 July; accepted 28 July 1988.

- Fruchter, A. S., Stinebring, D. R. & Taylor, J. H. *Nature* **333**, 237–239 (1988).
- Phinney, E. S., Evans, C. R., Blandford, R. D. & Kulkarni, S. R. *Nature* **333**, 832–834 (1988).
- Van den Heuvel, E. P. J. & van Paradijs, J. *Nature* **334**, 227–228 (1988).
- Kluźniak, W., Ruderman, M., Shaham, J. & Tavani, M. *Nature* **334**, 225–227 (1988).
- Ruderman, M., Shaham, J. & Tavani, M. *Astrophys. J.* (in the press).
- Fruchter, A. S. *et al. IAU Circular No.* 4617 (1988).
- Allington-Smith, H., Breare, M., Ellis, M., Gray, P. & Worswich, S. *WHT FOS User Manual* (Royal Greenwich Observatory, 1988).
- Oke, J. B. *Astrophys. J. Suppl.* **27**, 21–35 (1974).
- Stetson, P. B. *Publ. astr. Soc. Pacif.* **99**, 191–222 (1987).
- Manchester, R. N. & Taylor, J. H. *Astr. J.* **86**, 1953–1973 (1981).
- Lyne, A. G., Manchester, R. N. & Taylor, J. H. *Mon. Not. R. astr. Soc.* **213**, 613–639 (1985).
- Neckel, Th. & Klare, G. *Astr. Astrophys. Suppl.* **42**, 251–281 (1981).
- Weaver, H. & Williams, D. R. W. *Astr. Astrophys. Suppl.* **17**, 1–249 (1975).
- Dame, T. M. *et al. Astrophys. J.* **322**, 706–720 (1987).
- Savage, B. D. & Mathis, J. S. *A. Rev. Astr. Astrophys.* **17**, 73–111 (1979).
- Popper, D. A. *Rev. Astr. Astrophys.* **18**, 115–164 (1980).
- Van Paradijs, J. in *Accretion-driven Stellar X-ray Sources* (ed. Lewin, W. H. G. & van den Heuvel, E. P. J.) 189–260 (Cambridge University Press, 1983).
- Zapolsky, H. S. & Salpeter, E. E. *Astrophys. J.* **158**, 809–812 (1969).
- Backer, D. C., Kulkarni, S. R., Heiles, C., Davis, M. M. & Goss, W. M. *Nature* **300**, 615–618 (1982).
- Kippenhahn, R. & Thomas, H.-C. *Astr. Astrophys.* **75**, 281–290 (1979).
- Kuiper, L., van Paradijs, J. & van der Klis, M. *Astr. Astrophys.* (in the press).

## Optical detection and characterization of the eclipsing pulsar's companion

A. S. Fruchter\*, J. E. Gunn†, T. R. Lauer† & A. Dressler‡

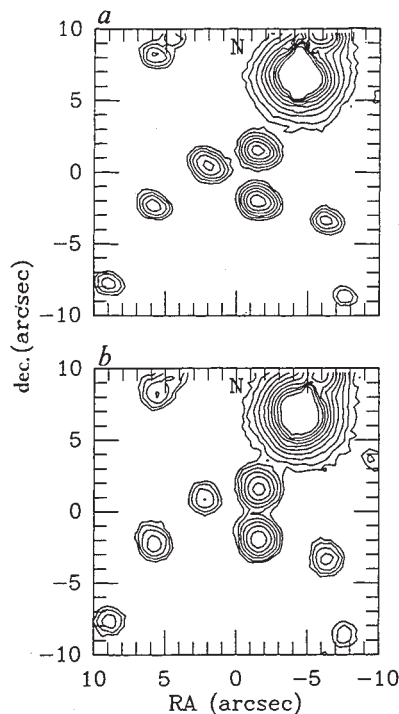
\* Joseph Henry Laboratories and Physics Department, Princeton University, Princeton, New Jersey 08544, USA

† Princeton University Observatory, Peyton Hall, Princeton, New Jersey 08544, USA

‡ Mount Wilson and Las Campanas Observatories, 813 Santa Barbara Street, Pasadena, California 91101, USA

The recently discovered millisecond pulsar PSR1957+20 is eclipsed for  $\sim 50$  min out of each 9.16-h orbit<sup>1</sup>; several authors suggested<sup>1–4</sup> that spin-down radiation from the pulsar is evaporating material from the surface of the  $\sim 0.02 M_\odot$  companion. Here we report observations intended to find the optical counterpart of the pulsar and to study luminosity variations with orbital phase. We show that the candidate optical counterpart<sup>5</sup> is actually two stars, one the true counterpart, the other an unrelated background star. The luminosity of the system varies by at least a factor of five, with a maximum of magnitude 20.3 when the pulsar hides the companion. The phase of the variations implies that the light comes mostly from the companion star, and suggests that it is tidally locked, with the bright side constantly illuminated by the pulsar. The companion has a low colour temperature,  $\sim 5,500 \text{ K}$ , which with its magnitude and the pulsar's dispersion-measure distance indicates a radius of  $0.15 R_\odot$ , about the size of a  $0.02 M_\odot$  hydrogen white dwarf.





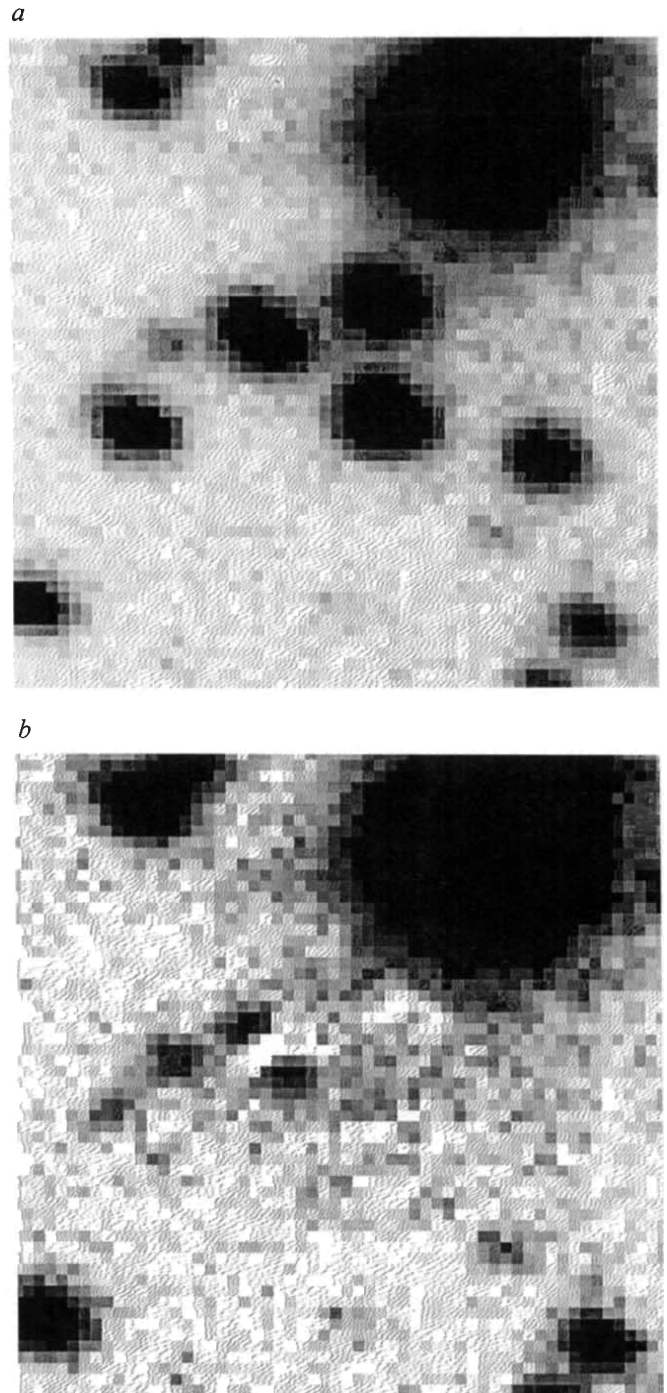
**Fig. 1** Contour plots of *i* images taken at pulsar anti-eclipse (*a*) and eclipse (*b*). Star X, the candidate object of Kulkarni *et al.*<sup>5</sup>, is the star to the left of centre. The candidate is substantially brighter at anti-eclipse than at eclipse. It is also noticeably extended towards the south-west at anti-eclipse, implying that the object is composed of more than one star.

During the nights of 9–11 May 1988 we obtained *g*, *r* and *i* band images of Star X, the suspected optical counterpart<sup>5</sup>, using the 4-shooter<sup>6</sup> on the 200-inch Hale Telescope. Images in all three filters were taken at orbital phases 0.79, 0.01 and 0.22, corresponding approximately to pulsar anti-eclipse, quadrature and pulsar eclipse, respectively. (We use the phase convention of Fruchter *et al.*<sup>1</sup> in which phase 0.0 is the ascending node of the pulsar's orbit.) One *r* image and a 600 s low-resolution spectrum were also obtained at orbital phase 0.6. The observing conditions were generally very good, but on the first night rapid changes in the temperature of the primary mirror caused the image quality to be degraded to nearly 2 inches.

Figure 1 shows contour plots of *i* band images taken at pulsar anti-eclipse and eclipse. Star X varied over the orbit in both strength and shape. At its brightest, during pulsar anti-eclipse, the image became elongated towards the south-west and its centroid moved by more than one pixel (1 pixel = 0.33 arcsec) from its position at eclipse, suggesting strongly that Star X was the sum of two objects.

Using the VISTA image-processing system<sup>7</sup>, we fitted point spread functions to Star X and five nearby stars. All stars other than Star X were stable relative to one another within statistical errors (1%), and were generally well fitted by point spread functions, but attempts to fit Star X with a single point spread function except at pulsar eclipse left residuals with a double humped profile (see Fig. 2). At anti-eclipse the residuals contained about 10% of the light of Star X, clearly indicating that the object was not a single star. We estimate that the pulsar's optical counterpart and a background star are separated by about  $0.8 \pm 0.1$  arcsec at a position angle of about  $40^\circ$ .

The excellent fit of the point spread function to Star X during pulsar eclipse implies that the counterpart might be completely dark at minimum. We used this possibility to simplify the photometry. We found the position of Star X relative to other stars at minimum and subtracted a star at this position from the



**Fig. 2** The PSR1957+20 field is an *i* image before (*a*) and after (*b*) fitting Star X and four of its neighbours with a single point spread function. While the four other stars are fitted successfully, Star X leaves behind systematic residuals indicative of the presence of two stars. The contrast of the second image has been enhanced to emphasize the residuals.

non-eclipse images to leave behind an image of the pulsar's optical counterpart. The centroid of the counterpart was then found and added to a list of stars containing the background component of Star X as well as five other nearby stars. Point spread functions were then fitted to all stars on the list, using the original unsubtracted image. A quadratic form was fitted simultaneously to the sky background, excluding from the fit pixels contaminated by a nearby saturated star (Star A of ref. 1).

Although trial changes in the position and magnitude of the background star caused the optical counterpart to vary in both

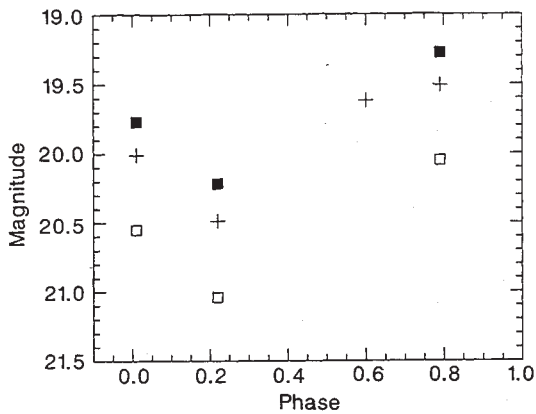


Fig. 3 A sparse light curve of Star X. Star X is the sum of two stars, one a background object and the other the optical counterpart of PSR1957+20. Filled boxes represent  $g$  magnitudes, crosses  $r$  magnitudes, and open boxes  $i$  magnitudes. The constancy of the colours is particularly striking. Errors in photometry are smaller than the symbols employed.

strength and location, the sum of the fits to the minimum light position and the optical counterpart (the total light in 'Star X') was found to be stable within statistical errors. Therefore we use this sum to measure the intensity of the pulsar system.

To obtain the absolute magnitude of the optical counterpart we observed, on 9 May, the standard star BD+17° 4708 (ref. 8) immediately following our observation of the PSR1957+20 system. Observations on successive nights were compared using the photometry of the five nearby stars (see Table 1 and Fig. 3). Applying the transformations of Kent<sup>9</sup>, our  $g$  and  $r$  band measurements at maximum imply a peak apparent visual magnitude  $M_V = 19.8$ .

We also took one 600 s low-resolution (20 Å) spectrum immediately after the  $r$  band image at orbital phase 0.6. The data, however, show only a faint continuum and no evidence of emission lines.

A surprising result of the photometry is the near constancy of the colours, implying both that the background star is nearly the same colour as the optical counterpart and that the PSR1957+20 system does not change colour substantially as a function of orbital phase. The data also show that the majority of the optical radiation cannot be from the pulsar, for if it were, the light curve would be flat away from radio eclipse. We are thus seeing the  $\sim 0.02M_\odot$  companion. Also, as the system is brightest when the pulsar is between the earth and the companion, and falls about half way to its minimum value at quadrature, the data agree with the hypothesis that the rotation of the companion is locked to the orbital period and that the companion therefore possesses a hot side facing the pulsar. Tests done on the images to determine whether the optical counterpart might continue to shine at minimum imply that less than half the flux from Star X at pulsar eclipse is due to the optical counterpart of PSR1957+20. Depending on exactly how much of the light at minimum one attributes to the companion, the peak visual luminosity of the pulsar's companion is between 20.05 and 20.35.

In any event, the data require that the luminous areas of the companion be quite uniform in temperature. For instance if one models the temperature of the surface of the companion as proportional to  $\cos^{1/4}(\theta)$  on the side facing the pulsar and zero on the back side (thus assuming local equilibrium between incident pulsar spin-down energy and re-radiation), one predicts far more luminosity and colour variation than is observed, regardless of the assumed interstellar reddening.

The uncorrected colour temperature of Star X at maximum is about 4,250 K. However, stars near the pulsar and at its

Table 1 Photometry of Star X as a function of PSR1957+20 orbital phase

| Phase | $m_g$ | $m_r$ | $m_i$ | $g-r$ | $r-i$ |
|-------|-------|-------|-------|-------|-------|
| 0.01  | 20.55 | 20.01 | 19.77 | 0.54  | 0.24  |
| 0.22  | 21.04 | 20.49 | 20.22 | 0.54  | 0.27  |
| 0.60  |       | 19.62 |       |       |       |
| 0.79  | 20.05 | 19.51 | 19.28 | 0.54  | 0.24  |

Star X is the sum of two stars, one a background object and the other the optical counterpart of PSR1957. The optical counterpart produces less than one half of the optical flux at minimum. All images were taken within  $\Delta\phi = 0.02$  of the phase quoted. The photometric errors are about 0.02 magnitudes for filter intensities and 0.03 magnitudes for colours.

dispersion measure distance ( $\sim 0.8$  kpc)<sup>1,10</sup> generally suffer a visual extinction of 1–2 magnitudes<sup>11</sup>, with a small number having extinctions as high as 3 magnitudes. Similarly, converting the H I column density<sup>12</sup> in the direction of the pulsar ( $2.5 \times 10^{21} \text{ cm}^{-2}$ ) to a visual extinction<sup>13</sup> one finds  $A_V = 1.5$ , leading to a surface temperature of  $\sim 5,500$  K for the companion. At this temperature and extinction the colour temperature changes by about 10% for a change in extinction of 0.5 magnitudes, implying an uncertainty in the colour temperature of about 500 K.

The colours of the background star are almost identical to the companion's and therefore correspond to a late G or early K spectral class. Were it at a distance of 1 kpc, however, its absolute visual magnitude would be  $M_V \approx 9.0$ , which is far too faint for a star on the main sequence with these colours. Thus we believe it is not associated with the pulsar but is a main sequence dwarf at a distance of a few kpc.

Because the surface of the companion will be in equilibrium with the radiation from the pulsar, we can use the object's colour temperature to calculate the fraction of the pulsar spin-down energy directed at the companion that penetrates to its photosphere. Assuming that the companion has a hot hemisphere at a uniform temperature of 5,500 K facing the pulsar and a cold backside which does not radiate significantly, that the moment of inertia of the pulsar is  $1.4 \times 10^{45} \text{ g cm}^2$ , that the pulsar wind is isotropic and that the mass of the pulsar is  $1.4M_\odot$ , we find that the fraction of the pulsar spin-down energy directed at the companion and re-radiated in the optical,  $f_{\text{rad}}$ , is

$$f_{\text{rad}} = 2.8^{+0.2}_{-0.6} \times 10^{-2} \left( \frac{T_c}{5,500 \text{ K}} \right)^4 \left( \frac{\dot{P}}{1.0 \times 10^{-19}} \right)^{-1}$$

where  $T_c$  is the temperature of the companion,  $\dot{P}$  is the period derivative of the pulsar (for comparison, the period derivative of the millisecond pulsar PSR1937+21 is  $1.0 \times 10^{-19}$ ) and the stated errors reflect our ignorance of the companion's flux at minimum. Unless  $\dot{P}$  is unexpectedly low, very little of the pulsar spin-down energy directed at the companion is re-radiated. If there is no unexpected mechanism for removing energy from the surface of the companion (such as strong magnetic activity), most of the energy in pulsar wind must be absorbed or deflected before it can reach the surface of the companion.

From the surface temperature, we can also determine the size of the companion. We estimate that

$$R_c = (0.15 \pm 0.04) \left( \frac{D}{0.8 \text{ kpc}} \right) R_\odot$$

where  $R_c$  is the radius of the companion and  $D$  is the distance to the pulsar. The major sources of error in the coefficient are the uncertainties in the visual extinction and the companion's surface temperature and flux, although the dispersion measure distance to the pulsar (0.8 kpc) could be inaccurate by as much as 50%. Nonetheless, the radius of a degenerate star of mass  $0.022M_\odot$  is  $\sim 0.046(1 + X_c)^{5/3} R_\odot$ , where  $X_c$  is the hydrogen mass



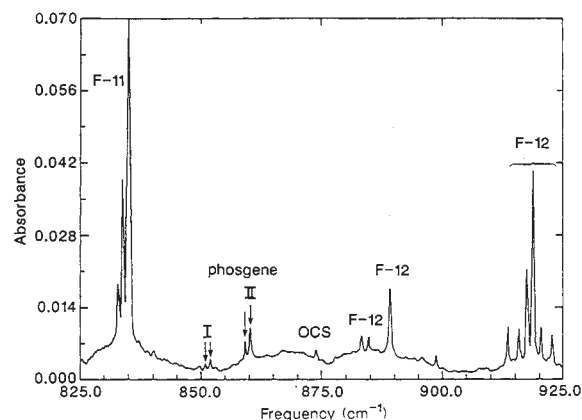
fraction<sup>14</sup>. For a primordial abundance of hydrogen, this is equal to 0.12R<sub>⊙</sub>, which is quite close to our best estimate of the size of the companion. The data therefore suggest that the companion does not have a large helium mass fraction.

The uncertainty in our conclusions could be reduced substantially by future optical observations aimed at measuring the extinction in this region of the sky, and by radio timing, which may soon allow the determination of the pulsar's period derivative and parallactic distance.

We thank G. Knapp for obtaining the estimate of the hydrogen column density in the direction of the pulsar and D. Schneider for the transfer functions of the Gunn filters. We also thank the staff of Palomar Observatory for their assistance. This work was funded in part by the NSF. A.S.F. is supported by a John von Neumann Fellowship.

Received 14 July; accepted 28 July 1988.

1. Fruchter, A. S., Stinebring, D. R. & Taylor, J. H. *Nature* **333**, 237–239 (1988).
2. Kluzniak, W., Ruderman, M., Shaham, J. & Tavani, M. *Nature* **334**, 225–227 (1988).
3. Phinney, E. S., Evans, C. R., Blandford, R. D. & Kulkarni, S. R. *Nature* **333**, 832–834 (1988).
4. Van den Heuvel, E. P. J. & van Paradijs, J. *Nature* **334**, 227–228 (1988).
5. Kulkarni, S. R., Djorgovski, S. G. & Fruchter, A. S. *Nature* **334**, 504–506 (1988).
6. Gunn, J. E. *et al. Optical Engineering* **26**, No. 8, 779–787 (1987).
7. Lauer, T. R., Stover, R. J. & Terndrup, D. M. *Lick Observatory tech. Rep. No. 34* (1983).
8. Oke, J. B. & Gunn, J. E. *Ap. J.* **266**, 713–717 (1983).
9. Kent, S. M. *Publ. astr. Soc. Pacif.* **97**, 165–174 (1985).
10. Lyne, A. G., Manchester, R. N. & Taylor, J. H. *Mon. Not. R. astr. Soc.* **213**, 613–639 (1983).
11. Neckel, Th. & Clare, G. *Astr. Astrophys. Suppl. Ser.* **42**, 251–281 (1980).
12. Weaver, H. & Williams, D. R. W. *Astr. Astrophys. Suppl.* **17**, 1–249 (1974).
13. Spitzer Jr, S. *Physical Processes in the Interstellar Medium* 156–157 (Wiley, New York, 1978).
14. Rappaport, S., Nelson, L., Ma, C. P. & Joss, P. C. *Astrophys J.* **322**, 842–851 (1987).



**Fig. 1** Infrared spectrum of phosgene in solid CO<sub>2</sub>, taken from a sample collected at ~13 km. Absorbances due to F-11 (CFCl<sub>3</sub>), F-12 (CF<sub>2</sub>Cl<sub>2</sub>), OCS and phosgene are marked. For phosgene in CO<sub>2</sub>, two different matrix environments are apparent (labelled I and II), giving rise to a doubling of the peaks in the spectrum. The absorption features shown here are due to the  $\nu_4$  vibration. Band assignments are made by a comparison with the known infrared spectrum of phosgene in an argon matrix<sup>20</sup>. The mixing ratio was determined by measurements of the heights of the absorption features from the two sites. The spectral intensities were calibrated by measuring the spectra of standard mixtures of phosgene in CO<sub>2</sub>, as previously described<sup>5</sup>.

## Phosgene measurements in the upper troposphere and lower stratosphere

S. R. Wilson, P. J. Crutzen, G. Schuster,  
D. W. T. Griffith\* & G. Helas

Max Planck Institute for Chemistry, PO Box 3060,  
D-6500 Mainz, FRG

\* Chemistry Department, University of Wollongong, PO Box 1144,  
Wollongong, New South Wales 2500, Australia

Industrial chlorofluorocarbons have now accumulated so much in the atmosphere that the ClO<sub>x</sub> radicals produced from their oxidation are causing substantial reductions in the ozone layer. Here we measure phosgene, which is one possible product from the oxidation of natural and industrial chlorinated hydrocarbons and which can oxidize further to form ClO<sub>x</sub>. Our measurements show a mixing ratio of 17 p.p.t.v. in the upper troposphere, and an average of 22 p.p.t.v. in the lower stratosphere. These values are substantially greater than those estimated with a model that only considers the photochemical breakdown of CCl<sub>4</sub>, indicating the possible significance of other more reactive chlorocarbon compounds, especially CHCl<sub>3</sub>, CH<sub>3</sub>CCl<sub>3</sub>, C<sub>2</sub>HCl<sub>3</sub> and C<sub>2</sub>Cl<sub>4</sub> and their oxidation products in supplying chlorine to the lower stratosphere.

Phosgene (COCl<sub>2</sub>) has been detected at the Earth's surface, in marine, continental and urban air masses in California<sup>1,2</sup>. Concentrations vary considerably, but show an average mixing ratio of ~25 p.p.t.v. (where p.p.t.v. is 1 in 10<sup>12</sup> molecules). Its presence at ground level has been explained in terms of the oxidation of chlorinated hydrocarbons such as the chloroethenes.

The atmospheric mixing ratio for the carbonyl halides phosgene, carbonyl difluoride (COF<sub>2</sub>) and COClF has been calculated in a one-dimensional model of the atmosphere, assuming that they are the products of oxidative chlorofluorocarbon (CFC) photolysis<sup>3</sup>. For phosgene, it was assumed that its forma-

tion follows the photochemical decay of CCl<sub>4</sub>. Using mixing ratios for the precursors relevant to 1976, a mixing ratio profile was calculated that increased rapidly from less than 1 p.p.t.v. at the tropopause, reaching a maximum of 15 p.p.t.v. at 25 km. This calculation did not consider the possibility of a tropospheric source for phosgene arising from OH-initiated oxidation of the chlorinated alkenes and alkanes. Of the above-mentioned carbonyl halides, only carbonyl difluoride has been previously observed in the stratosphere<sup>4</sup>.

The measurements reported here have been made using a matrix isolation-Fourier transform infrared spectroscopy technique<sup>5</sup>: the technique involves the cryogenic trapping of air samples at a temperature of 87 K, so that the material frozen out is primarily CO<sub>2</sub> and water. Following appropriate handling, the CO<sub>2</sub> forms a glassy matrix, with the condensable trace gases embedded in it. The infrared spectrum of this matrix can then be measured; the absorption peaks are characteristic of the specific trace gases and the height or area under each peak gives a quantitative measure of the amount of the trace gas present.

Figure 1 shows a portion of an infrared spectrum from such an air sample, with the absorption features that have been assigned to phosgene clearly visible. The spectral features were identified and the observed intensities calibrated by comparison with the spectra of mixtures of phosgene in CO<sub>2</sub> made using standard dilution techniques.

For the measurements described, the samples were collected aboard a Lear Jet (25D) on several flights between Germany and Spitzbergen (50° N–78° N). The date of each measurement flight is given in Table 1. In addition to the measurements of phosgene reported here, the mixing ratio of a number of other compounds were measured, including N<sub>2</sub>O, OCS, SO<sub>2</sub>, CFCl<sub>3</sub>, CF<sub>2</sub>Cl<sub>2</sub>, CHClF<sub>2</sub> and CH<sub>3</sub>CCl<sub>3</sub> (S.R.W., P.J.C., G.S., D.W.T.G. and G.H., manuscript in preparation).

The measured phosgene mixing ratios are shown in Fig. 2, plotted versus height relative to the calculated tropopause level. The tropopause height has been determined from the combination of the measured O<sub>3</sub> mixing ratios and the radiosonde data. The variation in the measurements in the stratosphere is partially due to meteorological differences in the air masses sampled.

**Table 1** Mixing ratios of phosgene in the upper troposphere and lower stratosphere

| Flight      | Tropopause height (km) | Date    | Troposphere (p.p.t.v.) | Stratosphere (p.p.t.v.) |
|-------------|------------------------|---------|------------------------|-------------------------|
| LJ6         | 12.1                   | 4/12/86 | 17.0 ± 2.7 (10)        | 23.4, 26.4, 24.6        |
| LJ7         | 9.9                    | 20/1/87 | 19.7, 20.1             | 22.7 ± 2.5 (7)          |
| LJ8         | 7.3                    | 21/2/87 | 14.8, 15.5, 17.2       | 23.8 ± 3.1 (6)          |
| LJ9         | 9.5                    | 22/4/87 | 15.0, 15.7             | 21.0 ± 2.1 (7)          |
| LJ10        | 9.4                    | 11/6/87 | 18.5 ± 2.1 (4)         | 20.4 ± 1.3 (7)          |
| All flights |                        |         | 17.3 ± 2.4 (21)        | 22.2 ± 2.6 (30)         |

A standard deviation for the distribution is given where enough measurements are available. For measurements within 0.5 km of the tropopause (4 values), the assignment to the troposphere or stratosphere was made on the basis of the ozone mixing ratios (troposphere < 100 p.p.b.v. ozone). The numbers in brackets after the mixing ratio are the number of measurements averaged.

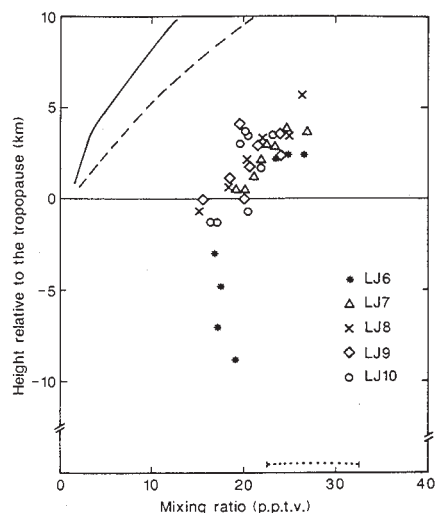
Mean values of the mixing ratio observed for phosgene for each flight are shown in Table 1. In the troposphere the average mixing ratio over all flights (21 measurements) is  $17.3 \pm 2.4$  p.p.t.v. In the stratosphere the average is  $22.2 \pm 2.6$  p.p.t.v. (30 measurements). No significant seasonal variation in the mixing ratio can be seen from these figures. The errors shown are the standard deviations of the measurements, and do not include an estimate of systematic errors, such as in the calibration. The absolute error in these measurements should be <20%.

The range of mixing ratios measured by Singh *et al.*<sup>2</sup> at ground level are also shown in Fig. 2. They suggest that the average mixing ratio for phosgene was 22 p.p.t.v. in 1976. If this is still the case today, there is a small decrease in mixing ratio between ground level and the upper troposphere. This is not surprising, because the vertical movement will often involve transport through clouds which should be scavengers for phosgene, but it is clear that more measurements are needed.

The mixing ratio profile of Crutzen *et al.*<sup>3</sup>, calculated by considering the oxidation of  $\text{CCl}_4$ , is also shown in Fig. 2. The mixing ratios of  $\text{CCl}_4$  used in this calculation are too small for the present day atmosphere, however. Therefore to allow comparison, both the calculated value and the calculation scaled by the amount of  $\text{CCl}_4$  measured<sup>6</sup> in the atmosphere is included. It can be seen that this underestimates the amount of phosgene significantly in both the stratosphere and the troposphere.

The most striking difference between the model calculation and the measured profile is that the tropospheric mixing ratios are comparable with those measured in the lower stratosphere, implying a significant source for phosgene in the troposphere. The primary route for the formation of phosgene is likely to be the oxidation of chlorinated ethenes<sup>1</sup> and the chlorinated alkanes methylchloroform, chloroform and possibly dichloromethane. Trichloroethene, tetrachloroethene and methylchloroform are produced in roughly equal quantities ( $\sim 5 \times 10^9$  mol  $\text{yr}^{-1}$ )<sup>7,8</sup>. Chloroform is mainly released from the sea, and can be assumed to have a similar source strength<sup>7</sup>. For the chlorinated alkenes, the dominant oxidation pathway through OH (ref. 10) produces phosgene, although not with unit reaction yield<sup>11</sup>. With these source strengths, an estimate of the lifetime of phosgene in the troposphere of about two months can be made, assuming steady state and unit reaction yields.

The primary sink in the troposphere is assumed to be liquid phase hydrolysis and heterogeneous decomposition<sup>1</sup>. The hydrolysis rate of phosgene in water has been measured<sup>12</sup>, but not under conditions relevant for the atmosphere. Further, to properly understand such processes, it is also necessary to consider in detail the dynamics of transport in the troposphere, in particular the passage of air through clouds. It is clear from our results that hydrolysis cannot be so fast as to stop transport of phosgene in the troposphere and into the stratosphere.



**Fig. 2** The measured mixing ratio of phosgene plotted against height relative to the tropopause. Where parallel measurements have been made the values have been averaged. The estimated uncertainty in the determination of the tropopause is  $\pm 0.5$  km. The solid line shows the mixing ratio of phosgene calculated by Crutzen *et al.*<sup>3</sup>, and the dotted line shows this calculation scaled by the known mixing ratio of the assumed precursor,  $\text{CCl}_4$ . At the bottom of the figure the dotted line indicates the range of mean mixing ratios measured at ground level by Singh *et al.*<sup>2</sup>.

The excess phosgene in the stratosphere in comparison to the calculated values can arise from either other sources of phosgene and/or a slower loss rate of phosgene than had been considered in this calculation. The other source that is known to be present in the stratosphere in significant amounts is methylchloroform ( $\text{CCl}_3\text{CH}_3$ ). Oxidation by OH is rather slow, but results in the formation of chloral ( $\text{CCl}_3\text{CHO}$ ) (ref. 13), which should then photolyse to form phosgene<sup>13,14</sup>. Measurements of the photolysis of methylchloroform show that it will not form phosgene<sup>13</sup>. Clearly both of these processes should be included in model calculations. It should be noted that the two most abundant chlorofluorocarbons, F-11 and F-12, do not produce phosgene.

Two loss processes are likely to be important in the stratosphere, photolysis and transport into the troposphere. There is at present no reason to doubt the photolysis rate used, but the significant tropospheric mixing ratio of phosgene will be important in the loss rate to the troposphere. The gradient in the phosgene mixing ratios indicates that the net flux is from the stratosphere to the troposphere, but this flux will be considerably smaller than would be the case if there were not significant amounts of phosgene in the troposphere. This implies that the chlorinated hydrocarbons that oxidize in the troposphere may significantly influence the phosgene mixing ratio in the stratosphere.

Our finding also raises the question of whether other relatively stable chlorine-containing oxidation products are present in the troposphere. For instance, the formation of  $\text{CCl}_3\text{O}_2\text{NO}_2$  during the oxidation of methylchloroform or chloroform ( $\text{CHCl}_3$ ) through the  $\text{CCl}_3$  radical has been inferred from laboratory measurements<sup>15</sup>. Although unstable at room temperature, below 260 K it is thermally more stable than  $\text{HNO}_4$  (ref. 16), which is present in the atmosphere in relatively large amounts<sup>17</sup>. Measurements of the absorption spectrum suggest that it is photochemically similar to  $\text{HNO}_4$  (ref. 18). On decomposition it forms phosgene. The chlorinated analogue of peroxyacetyl-nitrate,  $\text{CCl}_3\text{C}(\text{O})\text{O}_2\text{NO}_2$ , could also exist because chloral ( $\text{CCl}_3\text{CHO}$ ) is also an expected product during the oxidation of methylchloroform, but is probably photolysed before it can react further<sup>14</sup>. In the low- $\text{NO}_x$  background, atmosphere gases



like  $\text{CCl}_3\text{OOH}$  and  $\text{CCl}_3\text{CH}_2\text{OOH}$  will be formed. Such species could be important in the chlorine chemistry of the atmosphere as they can also provide a method by which chlorine could be transported into the stratosphere. Kinetic information on the chemistry of such intermediates is, however, very incomplete and atmospheric observations have not been made.

In the upper troposphere and lower stratosphere, phosgene will be important as a relatively stable reservoir species for chlorine. The decay of phosgene in the stratosphere can be expected to be mainly through photolysis, ultimately leading to  $2\text{ClO}_x$  and  $\text{CO}_2$ . Measurements of the ultraviolet spectrum<sup>19</sup> show that this photolysis will occur at wavelengths that are longer than those at which the chlorofluorocarbons dissociate, which implies that the photo-oxidation of phosgene can take place lower down in the stratosphere. In the lower stratosphere the photolysis of phosgene (and the other intermediate gases mentioned above) is probably the largest and a significant source of  $\text{ClO}_x$ . Such processes should be taken into account in future considerations on the effect of chlorinated hydrocarbons on lower stratospheric ozone.

We thank Flight Captains Gärtner and Lechner of the FVG for their cooperation during these flights. This work was par-

tially supported by the Bundesministerium für Forschung und Technologie.

Received 8 February; accepted 13 July 1988.

1. Singh, H. B. *Nature* **264**, 428–429 (1976).
2. Singh, H. B., Salas, L., Shigeishi, H. & Crawford, A. *Atmos. Envir.* **11**, 819–828 (1977).
3. Crutzen, P. J., Isaksen, I. S. A. & McAfee, J. R. *J. geophys. Res.* **83**, 345–363 (1978).
4. Rinsland, C. P. *et al. Geophys. Res. Lett.* **13**, 769–772 (1986).
5. Griffith, D. W. T. & Schuster, G. J. *atmos. Chem.* **5**, 59–81 (1987).
6. *Atmospheric Ozone, 1985* (World Meteorological Organization Report No. 16, Geneva, Switzerland, 1985).
7. Gidel, L. T., Crutzen, P. J. & Fishman, J. *J. geophys. Res.* **88**, 6622–6640 (1983).
8. Hammit, J. K. *et al. Nature* **320**, 711–716 (1987).
9. Rasmussen, R. A., Khalil, M. A. K. & Crawford, A. J. *Geophys. Res. Lett.* **9**, 704–707 (1982).
10. Atkinson, R., Aschmann, S. M., Fitz, D. R., Winer, A. M. & Pitts, J. N. Jr *Int. J. chem. Kinetics* **14**, 13–18 (1982).
11. Atkinson, R. *Chem. Rev.* **86**, 69–201 (1986).
12. Manogue, W. H. & Pigford, R. L. *A.I.Ch.E. J.* **6**, 494–500 (1960).
13. Nelson, L., Treacy, J. J. & Sidebottom, H. W. In *Proc. 3rd European Symp. Physico-Chemical Behaviour of Atmospheric Pollutants* (eds Versino, B. & Angeletti, A.) 258–263 (Reidel, Dordrecht, 1984).
14. Ohta, T. & Mizoguchi, I. *Int. J. chem. Kinetics* **12**, 717–727 (1980).
15. Simonaitis, R. & Heicklen, J. *Chem. Phys. Lett.* **62**, 473–478 (1979).
16. Baldwin, A. C. & Golden, D. M. *J. Phys. Chem.* **82**, 644–647 (1978).
17. Rinsland, C. P. *et al. Geophys. Res. Lett.* **13**, 761–764 (1986).
18. Morel, O., Simonaitis, R. & Heicklen, J. *Chem. Phys. Lett.* **73**, 38–42 (1980).
19. Moule, D. C. & Foo, P. D. *J. chem. Phys.* **55**, 1262–1268 (1971).
20. Schnöckel, H. *J. molec. Struct.* **29**, 123–131 (1975).

## Evidence from angle-resolved resonant photoemission for oxygen-2p nature of the Fermi-liquid states in $\text{Bi}_2\text{CaSr}_2\text{Cu}_2\text{O}_8$

T. Takahashi\*, H. Matsuyama\*, H. Katayama-Yoshida\*, Y. Okabe\*, S. Hosoya†, K. Seki‡, H. Fujimoto§, M. Sato§ & H. Inokuchi§

\* Department of Physics, Tohoku University, Sendai 980, Japan

† Institute for Materials Research, Tohoku University, Sendai 980, Japan

‡ Department of Materials Science, Hiroshima University, Hiroshima 730, Japan

§ Institute for Molecular Science, Okazaki 444, Japan

It has been generally accepted that a strong on-site Coulomb repulsion of the Cu 3d electrons dominates the electronic structure of the high-transition-temperature (high- $T_c$ ) superconductors. The on-site Coulomb repulsion has been evaluated as 6–7 eV, comparable with the valence-band width<sup>1,2</sup>. This strong correlation is thought to cause the Cu 3d electrons to be localized as in a Mott insulator, and doped holes may be transferred to oxygen sites, as the charge transfer energy is small compared with the correlation energy. These doped holes yield a substantial density of states at the Fermi level, characteristic of metals. There has been great effort to find and characterize the electronic states at the Fermi level, because these states relate directly to the mechanism of the high- $T_c$  superconductivity by providing Cooper pairs below  $T_c$ . Here we report the first direct evidence for the dominant oxygen-2p nature of the Fermi-liquid state in the high- $T_c$  superconductor, obtained using the technique of angle-resolved resonant photoemission.

A single crystal of  $\text{Bi}_2\text{CaSr}_2\text{Cu}_2\text{O}_8$ , typically  $5 \times 5 \times 0.5$  mm, was grown by a KCl flux technique<sup>3</sup>, in which a stoichiometric melt of  $\text{Bi}_2\text{O}_3$ ,  $\text{SrCO}_3$ ,  $\text{CaCO}_3$  and  $\text{CuO}$  powders, of 99.9% purity, and KCl (~20 per cent by weight) were cooled from 890 °C to 850 °C at a rate of  $0.5^\circ\text{C h}^{-1}$  and then quenched to room temperature. The crystal has a smooth plane along which cleavage is easy. Single-crystallinity was confirmed by X-ray diffraction. A magnetic-susceptibility measurement showed that this crystal becomes superconducting at 85 K. Photoemission spectroscopy was performed with an angle-resolved photoemission spectrometer at the UVSOR, Institute for Molecular

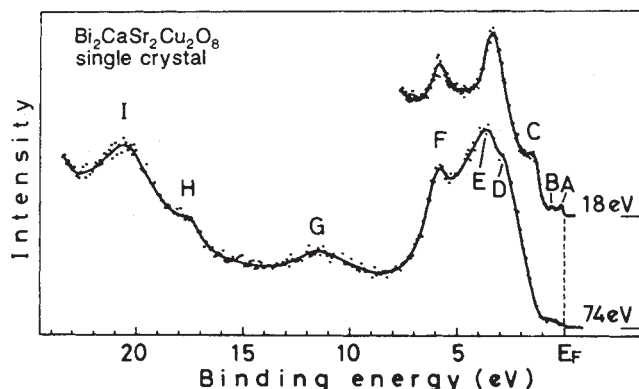
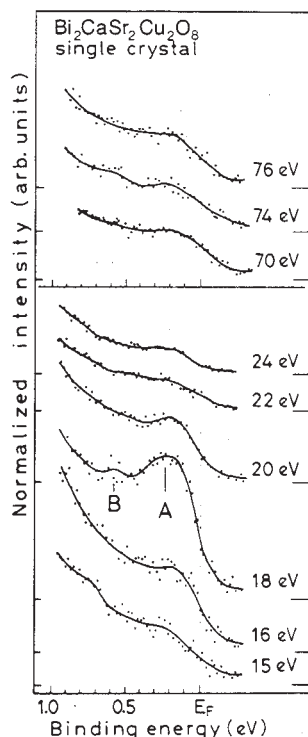


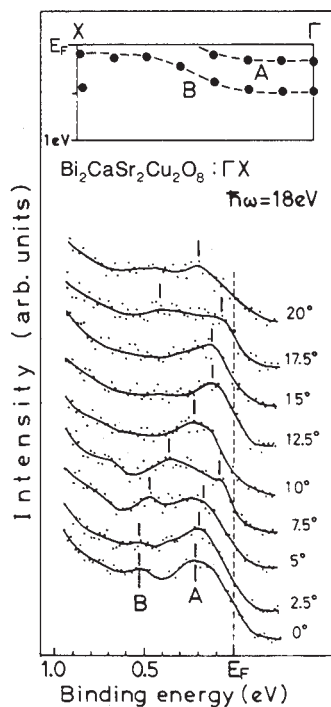
Fig. 1 Photoemission spectra of single-crystal  $\text{Bi}_2\text{CaSr}_2\text{Cu}_2\text{O}_8$  measured at  $\hbar\omega = 18$  and 74 eV in normal-emission arrangement.

Science, Japan. The energy resolution, estimated from the Fermi-edge of a gold film deposited in the spectrometer for calibration, was 0.2 eV, and the angular resolution was  $\sim 2^\circ$ . The single crystal was cleaved in the spectrometer to obtain a fresh surface and kept at room temperature during measurements. No surface degradation was detected throughout the measurement.

Figure 1 shows photoemission spectra of  $\text{Bi}_2\text{CaSr}_2\text{Cu}_2\text{O}_8$  measured at  $\hbar\omega = 18$  and 74 eV. We also studied resonant effects by varying the photon energy. Bands A–F (see Fig. 1) represent the valence band, which consists mainly of Cu 3d and O 2p states<sup>4–6</sup>. Bands H and I are ascribed to the O 2s and Sr 4p states respectively. The assignment of the O 2s band is confirmed by a resonant effect described below. Band G is a so-called two-hole bound state<sup>1,2</sup>, resulting from a many-body effect involving the Cu MVV Auger process. This band shows a strong enhancement at the Cu 3p core threshold ( $\hbar\omega = 74$  eV), reminiscent of La and Y systems<sup>1,2</sup>, indicating a strong electron correlation among d electrons. Bands D and F also show a resonance enhancement at  $\hbar\omega = 74$  eV, indicating a preferential contribution to the valence band from Cu 3d orbitals. The detailed interpretation of these resonant effects in the high-binding energy region will be described elsewhere. Here we concentrate on a small feature at the Fermi level (band A), because the electronic structure at the Fermi level plays a key role in high- $T_c$  superconductivity and is the focus of current controversy.



**Fig. 2** Photoemission spectra of single-crystal  $\text{Bi}_2\text{CaSr}_2\text{Cu}_2\text{O}_8$  in the vicinity of the Fermi level measured with photon energies near the O 2s (lower) and Cu 3p (upper) core thresholds in the normal-emission arrangement. The photon energy used is indicated on each spectrum. The intensity of photoemission is normalized to the incident-photon flux, independently for the O 2s and Cu 3p regions.



**Fig. 3** Angle-resolved photoemission spectra of single-crystal  $\text{Bi}_2\text{CaSr}_2\text{Cu}_2\text{O}_8$  measured in the high-symmetry direction  $\Gamma X$  for  $\hbar\omega = 18$  eV. Inset: band structure determined from angle-resolved photoemission.

Figure 2 shows the photon-energy dependence of the spectrum in the vicinity of the Fermi level for the energy regions near the O 2s and Cu 3p core thresholds. The spectra are normalized to the incident-photon flux. The photoemission intensity, and hence the density of states, at the Fermi level (band A) exhibits enhancement at the O 2s core threshold ( $\hbar\omega = 18$  eV) but no such enhancement is observed at the Cu 3p core threshold ( $\hbar\omega = 74$  eV), indicating, as explained below, that the electronic states at the Fermi level in  $\text{Bi}_2\text{CaSr}_2\text{Cu}_2\text{O}_8$  have a dominant O 2p nature.

Figure 3 shows angle-resolved photoemission spectra in the vicinity of the Fermi level measured in a high-symmetry direction in momentum space ( $\Gamma X$ ) for  $\hbar\omega = 18$  eV. The inset shows the band structure implied by the spectra. Comparison with a band structure calculation including the higher-binding-energy region has been described elsewhere<sup>7</sup>. Figure 3 shows that there are two bands, with a substantial energy dispersion of 0.2–0.5 eV, in the vicinity of the Fermi level and that one of them crosses the Fermi level midway between the  $\Gamma$  point and the Brillouin-zone boundary. This is clear evidence for the existence of a Fermi-liquid state in the high- $T_c$  superconductor and is consistent with the recently discovered BSC-like enhancement of the enriched  $^{17}\text{O}$  nuclear relaxation rate below  $T_c$  (ref. 8). Both experiments indicate that doped holes, residing mainly on the oxygen site owing to the strong on-site Coulomb repulsion at the copper site, behave like a Fermi liquid. The resonant effect observed at the O 2s core threshold (Fig. 2) is due to 2s–2p absorption, implying that there is an empty O 2p state just above the Fermi level. This supports recent results of electron-energy-loss-spectroscopy (EELS)<sup>9</sup>, in which a small absorption peak just above the O 1s absorption edge was ascribed to empty O 2p states.

Thus, the present experimental results provide a clear picture of the electronic structure of the high- $T_c$  superconductor, which consists of two major features: a Fermi-liquid state with predominantly O 2p character and strongly localized states away from the Fermi level with both Cu 3d and O 2p character. Superconductivity could be driven by Cooper-pairing of the O 2p holes in the Fermi-liquid state.

We thank Professors M. Tachiki and Y. Kuramoto for useful discussion. This work was supported by a grant from the Ministry of Education, Japan.

Received 15 July; accepted 29 July 1988.

1. Takahashi, T. *et al.* *Phys. Rev.* **B36**, 5686–5689 (1987).
2. Shen, Z. *et al.* *Phys. Rev.* **B36**, 8414–8428 (1987).
3. Katayama-Yoshida, H., Ikabe, Y., Takahashi, T. & Hosoya, S. *Jap. J. appl. Phys.* (submitted).
4. Hybertsen, M. S. & Mattheiss, L. F. *Phys. Rev. Lett.* **60**, 1661–1664 (1988).
5. Krakauer, H. & Pickett, W. E. *Phys. Rev. Lett.* **60**, 1665–1667 (1988).
6. Massidda, S., Yu, J. & Freeman, A. J. *Physica C52*, 251–258 (1988).
7. Takahashi, T. *et al.* *Phys. Rev. Lett.* (submitted).
8. Kitaoka, Y. *et al.* *Nature* (submitted).
9. Nücker, N., Fink, J., Fuggle, J. C., Durham, P. J. & Temmerman, W. M. *Phys. Rev.* **B37**, 5158–5163 (1988).

## Controlled crystallization of $\text{CaCO}_3$ under stearic acid monolayers

Stephen Mann\*, Brigid R. Heywood\*, Sundara Rajam\* & J. Derek Birchall†

\* School of Chemistry, University of Bath, Bath BA2 7AY, UK

† ICI Advanced Materials, Runcorn, Cheshire WA7 4QE, UK

A fundamental concept in the study of biomineralization concerns the molecular recognition of inorganic materials at organized organic macromolecular substrates<sup>1</sup>. Here we investigate this concept through the use of stearic acid monolayers in the controlled crystallization of  $\text{CaCO}_3$  from supersaturated solutions. Whereas crystallization in the absence of a monolayer results in rhombohedral calcite crystals, the presence of an organized monolayer gives



Table 1 XRD values

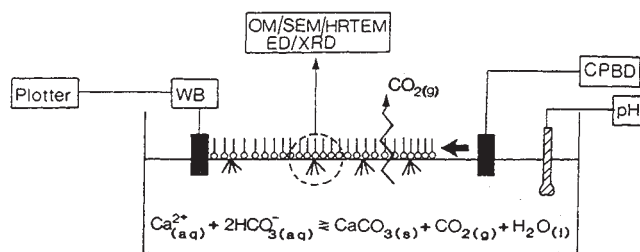
| Trough | Experimental Control | Theoretical Calcite | Vaterite |
|--------|----------------------|---------------------|----------|
| 4.241  | 3.843                | 3.852               | 4.245    |
| 3.576  |                      |                     | 3.577    |
| 3.290  | 3.030                | 3.030               | 3.296    |
|        | 2.831                | 2.834               |          |
| 2.726  |                      |                     | 2.735    |
|        | 2.262                | 2.284               |          |
| 2.202  |                      |                     | 2.219    |
| 2.112  |                      |                     | 2.122    |
|        | 2.079                | 2.094               |          |
| 2.060  |                      |                     | 2.065    |
|        | 1.901                | 1.907               |          |
|        | 1.865                | 1.872               |          |
| 1.857  |                      |                     | 1.857    |
| 1.820  |                      |                     | 1.825    |
| 1.645  |                      |                     | 1.648    |
|        | 1.619                | 1.604               |          |
|        | 1.591                | 1.582               |          |
| 1.534  |                      |                     | 1.544    |
| 1.471  |                      |                     | 1.477    |
|        | 1.468                | 1.506               |          |
|        | 1.437                | 1.440               |          |
|        | 1.417                | 1.416               |          |
| 1.358  |                      |                     | 1.351    |
|        | 1.334                | 1.336               |          |
| 1.308  |                      |                     | 1.331    |
| 1.283  |                      |                     | 1.286    |
|        | 1.175                | 1.177               |          |
|        | 1.141                | 1.141               |          |

All analyses were performed on a Phillips powder diffractometer—Cu K $\alpha$  radiation ( $\lambda = 1.5405 \text{ \AA}$ ).

rise to oriented vaterite formation. The vaterite nuclei are aligned with their (0001) face parallel to the plane of the organic substrate and develop initially in the form of disk-shaped single crystals. The degree of compression of the monolayer dictates the homogeneity of vaterite nucleation. In particular, partially compressed films are optimal for controlled crystallization, suggesting that the mobility of organic surfaces may be of general importance. Our results can be explained by electrostatic and stereochemical interactions at the inorganic–organic interface and these observations support current theories of biomineralization, as well as being of potential significance in the crystal engineering of microscopic inorganic assemblies<sup>2</sup>.

The influence of structured organic surfaces on the oriented overgrowth of inorganic minerals is a fundamental aspect of biomineralization. Addadi and co-workers<sup>3,4</sup> have studied the effect of  $\beta$ -pleated sheet proteins on the crystallization of calcite and showed that a cooperative effect arising from electrostatic, structural and stereochemical requirements gave rise to oriented calcite nucleation. Mann *et al.*<sup>5</sup> used organized phospholipid vesicles to control the structure of iron oxide phases formed from aqueous solution. Recently, Landau *et al.*<sup>6,7</sup> showed that oriented growth of glycine and, to a lesser extent, NaCl takes place at the air/water interface of Langmuir monolayers. In this paper, we describe the role of Langmuir monolayers in determining the structure, morphology and crystallographic orientation of CaCO<sub>3</sub> crystals formed from supersaturated solutions.

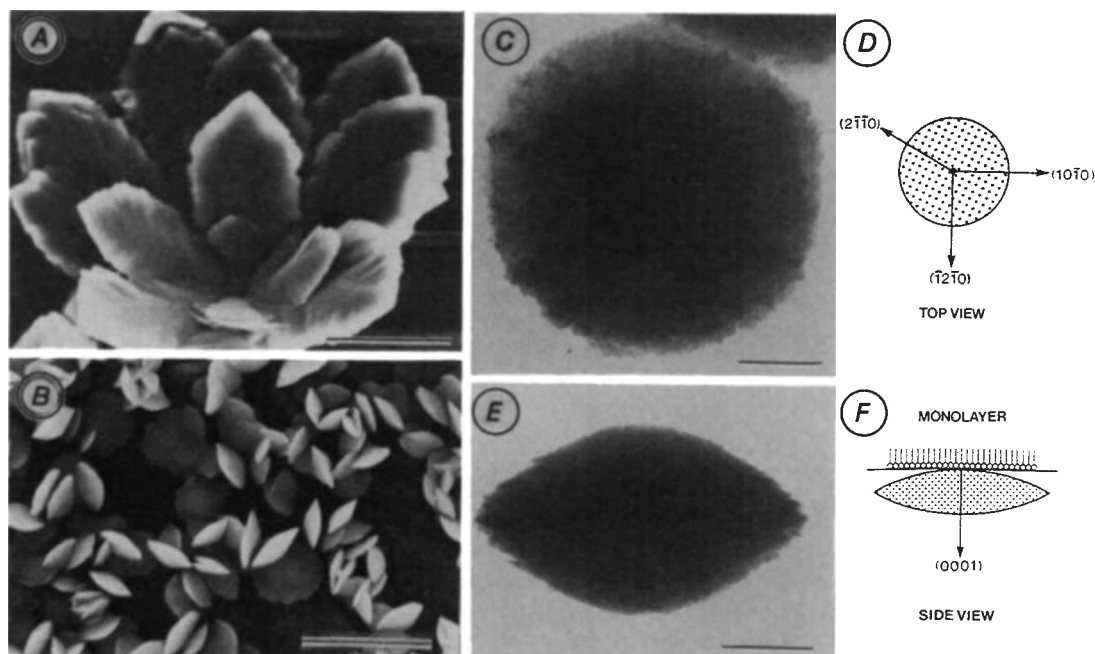
The experimental procedure for growing calcium carbonate crystals under stearic acid monolayers is shown diagrammatically in Fig. 1. In the absence of a monolayer, rhombohedral calcite crystals were formed (Table 1) at the surface, sides and bottom of the trough. Crystal growth at the solution surface was favoured as a result of continual CO<sub>2</sub> outgassing and the local supersaturation levels remained low, as shown by the absence



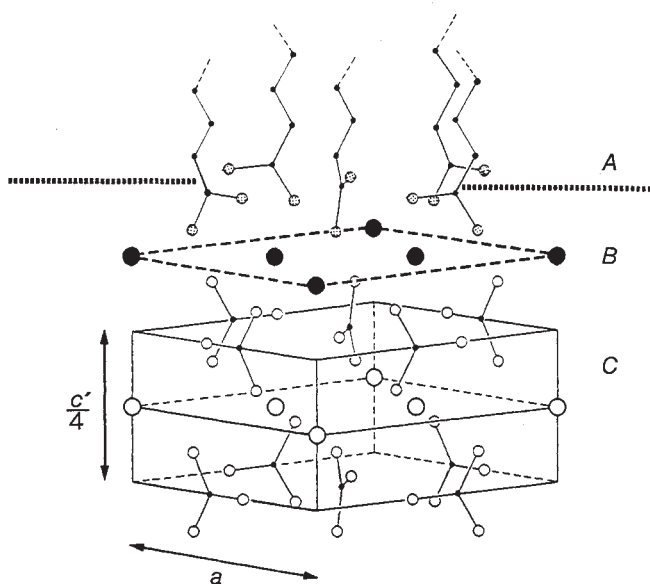
**Fig. 1** Experimental procedure for the growth of CaCO<sub>3</sub> crystals under stearic acid monolayers. Monolayer films of stearic acid (CH<sub>3</sub>(CH<sub>2</sub>)<sub>16</sub>COOH) were deposited on supersaturated calcium bicarbonate solutions freshly prepared using the method of Kitano<sup>10</sup>. A value of  $S = 10^3$ , where  $S$  is a measure of the maximum attainable supersaturation (namely that generated under conditions of instantaneous outgassing of all the CO<sub>2</sub>) was estimated by titrimetric analysis. The experimental pH (5.8–6.0) was above the  $pK_a$  (5.1) of stearic acid. Films were either uncompressed or compressed to liquid (20 mN m<sup>-1</sup>) or solid (40 mN m<sup>-1</sup>) phases and then left undisturbed for periods up to 16 h, during which crystal growth occurred by CO<sub>2</sub> outgassing. Crystals were collected from the monolayer surface at various times for structural and morphological analysis. Bulk samples for X-ray diffraction (XRD) were obtained by sweeping the surface with the barrier and collecting the aggregated crystals on glass slides. Crystals for optical microscopy (OM) were viewed *in situ* or on glass slides dipped through the films. Samples for scanning electron microscopy (SEM) were obtained by carefully touching the film surface with stainless steel stubs so that crystals were transferred by surface tension with only minimal disruption of their crystallographic organization. Crystals grown for 0–60 min were collected by gently dipping nitrocellulose-covered carbon-coated copper electron microscope grids once through the monolayer and were studied by high resolution transmission electron microscopy (HRTEM) and electron diffraction (ED). WB, Wilhelmy balance; CPBD, constant perimeter barrier device.

of kinetically stabilized polymorphs such as aragonite and vaterite. In contrast, in the presence of fully and partially compressed monolayers, crystallization was confined to the organic surface and the crystal structure was exclusively vaterite (Table 1). The mature vaterite crystals had a narrow particle-size distribution (mean 75  $\mu\text{m}$ ,  $\sigma = 10 \mu\text{m}$ ) and exhibited a complex pseudo-hexagonal floret morphology (Fig. 2A). The early crystals developed anisotropically along the organic interface in the form of lens-shaped disks (Fig. 2B). Disks of diameter  $\sim 3 \mu\text{m}$  were subsequently transformed into the floret habit by radial outgrowth into bulk solution. *In situ* optical microscopical investigations combined with scanning electron microscope observations indicated that the ventral surface of each disk was apposed to the monolayer with their radial axis normal to the plane of the interface. Electron diffraction patterns were recorded from individual disks and showed that the particles were single crystals of vaterite and that the radial axis corresponded to the crystallographic [0001] (*c*) axis (Fig. 2C and D). Viewed from the side, the disks were lens-shaped and gave diffraction patterns corresponding to either the crystallographic *a* ([1210]) axes or the [1010] axis (Fig. 2E and F). These results indicate that the crystallographic *c* and *a* axes were oriented perpendicular and parallel to the monolayer surface respectively and that nucleation of the disks occurred on the (0001) face of vaterite.

Compression isotherms (data not shown) for stearic acid on pure water gave a value for the limiting area per molecule of 22  $\text{\AA}^2$ . Films compressed on calcium-containing solutions gave a slightly expanded limiting area of 24  $\text{\AA}^2$  and a characteristic reduction in the surface pressure during the liquid-to-solid phase transformation<sup>8</sup>. These results indicated that Ca-binding to the carboxylate headgroups preceded nucleation at the monolayer surface. The formation of a Stern layer of Ca counterions favours



**Fig. 2** *A*, SEM of mature vaterite floret. The florets exhibit a complex crystalline pseudo-hexagonal morphology. Plate-like outgrowths radiate from a central disk-shaped core. Scale bar, 10  $\mu\text{m}$ . *B*, SEM of vaterite disks. The early crystals develop anisotropically along the monolayer in the form of lens-shaped disks. Radial outgrowth into the solution subphase transforms the disks into the mature floral habit. Scale bar, 5.0  $\mu\text{m}$ . *C*, TEM of vaterite disk, top view. Scale bar, 0.5  $\mu\text{m}$ . *E*, TEM of vaterite disk, side view. Scale bar, 0.5  $\mu\text{m}$ . *D*, *F*, Relationship between the morphology of the growing crystals, their crystallographic orientation and the stearic acid monolayer. The crystallographic axes  $c$  (0001) and  $a$  ( $2\bar{1}10$ ;  $1\bar{2}10$ ;  $10\bar{1}0$ ) were determined from electron diffraction patterns of the disks.



**Fig. 3** Proposed organization of the interface between stearic acid monolayers and nascent vaterite nuclei. *A*, Monolayer surface showing carboxylates aligned perpendicular to the air/water interface; the minimum inter-headgroup spacing for stearic acid is  $\sim 4.6$  Å; *B*, Stern layer of headgroup-bound  $\text{Ca}^{2+}$  ions; the uncharged layer is electrostatically equivalent to the (0001) face of vaterite, although the Ca-Ca distances are different; *C*, vaterite sub-cell showing the stereochemical arrangement of carbonates perpendicular to the (0001) face;  $c'/4 = 4.25$  Å,  $a = 7.15$  Å. (●, ○) Ca atoms, Stern layer and vaterite cells respectively; (○, ●) O atoms, vaterite cell and carboxylates respectively; (●) C atoms.

the oriented nucleation of crystal faces consisting of only Ca atoms and the (0001) face of vaterite fits this criterion (Fig. 3). Charge accumulation, however, cannot account for the structural selectivity of vaterite, as the (0001) face of calcite is also uni-charged. The trigonal planar carbonate anions are oriented parallel to the (0001) face of calcite, whereas in vaterite they are aligned perpendicularly, and this latter arrangement is equivalent to the orientation of the carboxylate headgroups with respect to the Ca-bound Stern layer. Thus, the stereochemical arrangement of the carboxylates in conjunction with Ca binding generates a two-layer subunit cell motif of the vaterite structure, rather than calcite (Fig. 3). Subsequent addition of carbonate from solution will be similarly stereochemically directed extending the vaterite motif towards a stable nucleus.

We emphasize that the mechanism proposed above does not involve the epitaxial matching of lattice dimensions in the (0001) face and headgroup spacings in the compressed monolayer film. The inter-headgroup spacing on the fully compressed monolayer is  $\sim 4.6$  Å, compared with a Ca-Ca distance of 4.13 Å in vaterite (4.96 Å in calcite). Thus the Stern layer shown in Fig. 3 is expanded compared with the layer of Ca atoms in the vaterite unit cell. This point highlights the importance of stereochemical and electrostatic matching because these factors can override the structural mismatch at the interface. Moreover, if geometric matching was a fundamental property of the system, we would expect calcite to be favoured over vaterite because the former has Ca-Ca distances that can be spanned by the carboxylate headgroups, for example in a partially compressed film.

Whereas fully compressed monolayers promoted rapid vaterite nucleation indiscriminately across the entire interface, resulting in small disks with a wide distribution of diameters (mean 1.32  $\mu\text{m}$ ,  $\sigma = 0.47$   $\mu\text{m}$ ,  $t = 35$  min), and nucleation on uncompressed films was slow and comparable with calcite nucleation from bulk solution, nucleation on partially compressed monolayers was intermediate in rate, as shown by a



narrow size distribution of relatively large (mean  $2.23 \mu\text{m}$ ,  $\sigma = 0.18 \mu\text{m}$ ,  $t = 35 \text{ min}$ ) but fewer vaterite disks. We propose that the rate of nucleation on liquid-phase monolayers is optimal because it requires the Ca-induced local ordering of stearate molecules into a configuration tailored for nucleation. Once formed, these sites are autocatalytic and successfully compete against the formation of new nucleation centres on the monolayer such that nucleation is confined to a limited number of discrete sites across the organic surface.

Finally, we note that organized macromolecular surfaces such as  $\beta$ -pleated sheet proteins and phospholipid membranes are components of many biological systems involving calcification<sup>9</sup>. With regard to the above observations, optimal control of biomineralization could be achieved on surfaces that are activated for nucleation through specific conformational changes induced by ion-binding, rather than on surfaces sterically predisposed and hence highly catalytic for nucleation. A mechanism of this type would be synergic as the coordination of specific ions present under supersaturated concentrations induces local interfacial modifications that, in turn, direct the flow of additional ions to the site and the subsequent development of the inorganic nucleus.

We thank Dr N. H. C. Sparks for assistance with scanning electron microscopy. The work was supported in part by SERC. S.R. is in receipt of a SERC/CASE studentship with ICI plc.

Received 27 April; accepted 19 July 1988.

1. Mann, S. *Nature* **332**, 119–124 (1988).
2. Calvert, P. D. & Mann, S. *J. Mater. Sci.* (in the press).
3. Addadi, L. & Weiner, S. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4110–4114 (1985).
4. Addadi, L., Moradian, J., Shay, E., Maroudas, N. G. & Weiner, S. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2732–2736 (1987).
5. Mann, S., Hannington, J. P. & Williams, R. J. P. *Nature* **324**, 565–567 (1986).
6. Landau, E. M., Levanon, M., Leiserowitz, L., Lahav, M. & Sagiv, J. *Nature* **318**, 353–356 (1985).
7. Landau, E. M. *et al. Molec. Cryst. liq. Cryst.* **134**, 323–335 (1986).
8. Gordziel, M. J., Flanagan, D. R. & Swarbrick, J. J. *Colloid. Interface Sci.* **86**, 178–184 (1981).
9. Weiner, S. *CRC Crit. Rev. Biochem.* **20**, 365–408 (1986).
10. Kitano, Y. *Bull. chem. Soc. Japan* **12**, 1980–1985 (1962).

## Observed velocity fluctuations on a major Antarctic ice stream

S. N. Stephenson\* & R. A. Bindshadler†

\* Science Applications Research, 4400 Forbes Boulevard, Lanham, Maryland 20706, USA

† Code 671, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

Most of the Antarctic ice sheet is drained by systems comprising an area of slow sheet flow which converges into a much faster valley glacier or ice stream. On reaching the coast, this ice then either enters an ice shelf confined by embayments and ice rises, or calves off into the ocean. The higher speed of ice streams enables them to respond to changes in their environment faster than the slower ice sheet. In the West Antarctic, where the ice sheet moves across bedrock well below sea level, the ice streams have the potential to rapidly increase or decrease the ice discharge into the sea<sup>1</sup>. The timescale for these changes has been placed at a few hundred years. If this is true, then changes should be detectable by careful measurements on a decadal timescale. By comparing recent ice velocity measurements with those collected ten years ago, we establish that the ice in the mouth of Ice Stream B has decelerated by ~20%. We discuss the possible causes of this deceleration on the basis of our knowledge of the current regional dynamics, and the possible ramifications of this deceleration on future ice stream behaviour.

Doppler satellite tracking methods were used to obtain ice velocities during the Ross Ice Shelf Geophysical and Glaciologi-

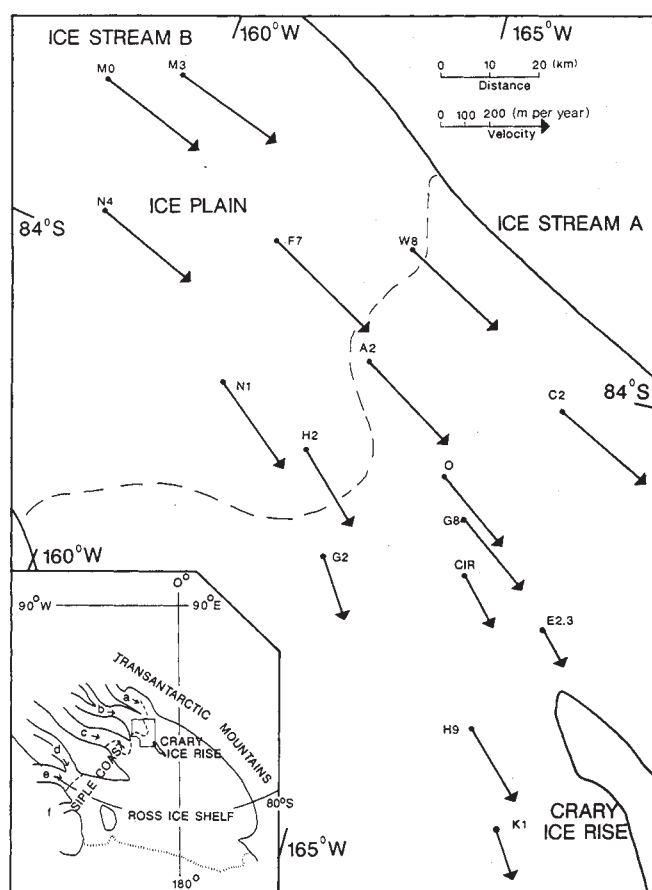


Fig. 1 Ice velocities upstream of Crary Ice Rise and on the ice plain of Ice Stream B. The thick continuous lines indicate ice-stream boundaries (after ref. 8, with modifications). The broken line indicates the grounding line. The inset shows the relative positions of the ice streams feeding the Ross Ice Shelf.

cal Survey (RIGGS) between 1972 and 1974, and then again during the current Siple Coast Program (SCP) begun in 1983. Both programs collected data on a 100-km section of ice extending from the ice shelf just upstream of Crary Ice Rise on to the ice plain of Ice Stream B (Fig. 1). The term ice plain refers to the area immediately upstream of the grounding line which has shallow surface slope similar to an ice shelf, and experiences only a very small amount of bed friction<sup>2</sup>. Velocities were measured by determining the position of markers twice, taking the velocity as being the distance between positions divided by the time interval between measurements. The velocities are presented in Table 1.

The velocities during RIGGS were determined using precise satellite ephemerides, but data from only a few (~5) satellite passes. The accuracy of these measurements is discussed by Thomas *et al.*<sup>3</sup>. They do not give details of the individual determinations at each site, but estimate an error of 15 m per yr for all the velocities. They regard this error as a conservative estimate because a systematic part of the positional error does not contribute to the velocity.

The velocities determined during SCP used the broadcast ephemeris and a simple model to compensate for tropospheric refraction. Data from a minimum of 5 and a maximum of 28 passes were used in the velocity determinations which are compared with RIGGS data. MacDonald and Whillans<sup>4</sup> estimate that the error in the isolated point-position determination is near 30 m, but they do not estimate how much of this error is systematic with position. Any systematic error would not contribute to the error in velocity. Four velocities at nearby stations,

**Table 1** Observed velocity data

| Station name | Latitude (deg. min s) | Longitude (deg. min s) | Velocity Magnitude (azimuth) |              |
|--------------|-----------------------|------------------------|------------------------------|--------------|
|              |                       |                        | (m per yr)                   | w.r.t. North |
| RIGGS data   |                       |                        | 1972-74                      |              |
| F7           | 84 07 11              | 162 03 52              | 530 (296)                    |              |
| G8           | 83 43 15              | 166 31 15              | 384 (306)                    |              |
| H9           | 83 20 51              | 167 25 27              | 348 (316)                    |              |
| SCP data     |                       |                        | 1984-85                      | 1985-87      |
| A2           | 83 57 29              | 164 16 24              | 459 (300)                    |              |
| C1R          | 83 37 14              | 166 44 31              | 245 (318)                    | 247 (319)    |
| C2           | 83 57 17              | 167 45 17              | 458 (296)                    | 452 (298)    |
| E2.3         | 83 33 09              | 168 12 42              | 172 (319)                    | 171 (312)    |
| G2           | 83 35 23              | 164 25 00              | 283 (326)                    |              |
| H2           | 83 46 09              | 163 40 14              | 362 (311)                    |              |
| K1           | 83 10 24              | 168 09 10              | 225 (326)                    | 229 (331)    |
| O            | 83 47 14              | 166 01 28              | 369 (302)                    | 376 (306)    |
| MO           | 84 17 42              | 158 10 27              | 471 (285)                    |              |
| M3           | 84 21 13              | 159 26 11              | 470 (285)                    |              |
| N1           | 83 50 30              | 161 56 00              | 402 (305)                    |              |
| N4           | 84 03 44              | 159 01 17              | 450 (288)                    |              |
| W8*          | 84 10 44              | 164 27 22              | 486 (297)                    |              |

\* First Doppler satellite tracking position from S. Shabtaie (personal communication).

not used in this study, were recalculated using precise ephemerides to estimate the error in determining the velocity using the broadcast ephemerides. The difference in velocity calculated using each method averaged 7 m per yr, with one velocity changing 14 m per yr. Additionally, when we compare the velocities determined over the period 1984 to 1985 with those that were also determined over the two-year interval 1985 to 1987 (see Table 1), we note that the average misclosure was 4 m per yr with a maximum of 7 m per yr. From Table 1 and other velocity comparisons not presented here, we conclude that the  $3\sigma$  error (99% confidence limit) in the velocities is about 25 m per yr. Given these estimates of the error, the combined  $3\sigma$  error in comparing RIGGS and SCP velocities is 51 m per yr. Thus we feel that if any velocity difference exceeds about 50 m per yr, then it is likely to be real.

To permit comparison of the SCP velocities with those of RIGGS, an interpolation scheme was developed where the

velocity at a point was determined by a linear interpolation between any three velocity measurements. The data used are given in Table 1. To test the accuracy of this method, control calculations were performed to predict the value of known SCP velocities from other SCP data, the results of which are shown in Table 2. All the azimuths agree well, as do the derived velocities for H2 and MO. But the misclosure in the velocity magnitude is significant for station O, and may be the result of local grounding between stations CIR and O.

Table 3 shows the comparison between RIGGS data and SCP-derived estimates. Table 3 indicates that both G8 and F7 have slowed by ~60 m per yr, or 17% and 11% respectively. The decrease in the velocity at H9 is about 100 m per yr. The position of H9 is close to an intense crevasse field<sup>5</sup> associated with the Crary Ice Rise. We expect that the strain field may vary considerably in the region, so we do not expect the same accuracy in predicting the velocity at this station.

The change in ice velocity over time is statistically significant. It is the first measured deceleration of an Antarctic ice stream, and it is remarkable because of its magnitude in ten years. The deceleration is not large enough to be detected by SCP measurements alone, that is by comparing velocities averaged from 1984 to 1985 with the velocities from 1985 to 1987 (see Table 1).

We now consider whether the observed slowing is just a local and perhaps short-lived phenomenon or whether it is related to changes throughout the whole drainage system or perhaps caused by external forcing.

Possible causes of the slowing could be the response of the ice stream to: (1) changes in the backforce exerted by Crary Ice Rise that were caused by changes in the boundary of the ice rise<sup>5</sup>. That is, the ice rise becomes more or less steamlined. (2) An increase in the thickness of the ice shelf surrounding Crary Ice Rise that would lead to a larger backforce on the ice stream<sup>6</sup>. Such an increase follows from the calculated positive mass balance in the region of Crary Ice Rise<sup>7</sup>. (3) Widening of Ice Stream B over time as a result of the observed decrease<sup>8</sup> in the discharge rate of Ice Stream C (Fig. 1 insert). (4) Changes in the ice-plain bed structure or hydraulics. (5) Advection of thicker ice as a result of the negative mass balance at the head of Ice Stream B (ref. 9). (6) The height above buoyancy increasing a few per cent as a result of isostatic rebound. Greischar and Bentley<sup>10</sup> estimate an uplift of 25 cm in 10 yr.

It is not possible to determine the quantitative effect of each of these changes without further study. Jezek<sup>11</sup> inferred the occurrence of transient variations in the ice-shelf flow, and our

**Table 2** Control estimates

| Name | Measured velocity    |         | Stations used to determine velocity field | Derived velocity     |         | Velocity difference  |         |
|------|----------------------|---------|-------------------------------------------|----------------------|---------|----------------------|---------|
|      | magnitude (m per yr) | azimuth |                                           | magnitude (m per yr) | azimuth | magnitude (m per yr) | azimuth |
| H2   | 362                  | 311     | A2, N1, G2                                | 361                  | 310     | -1                   | -1      |
| O    | 376                  | 306     | A2, H2, CIR                               | 340                  | 306     | -36                  | 0       |
| MO   | 471                  | 285     | A2, N4, M3                                | 465                  | 283     | -6                   | -2      |

**Table 3** Comparison of RIGGS and SCP velocities

| Name | Measured velocity    |         | Stations used to determine velocity field | Derived velocity     |         | Velocity difference  |         |
|------|----------------------|---------|-------------------------------------------|----------------------|---------|----------------------|---------|
|      | magnitude (m per yr) | azimuth |                                           | magnitude (m per yr) | azimuth | magnitude (m per yr) | azimuth |
| G8   | 384                  | 306     | O, CIR, E2.3                              | 326                  | 310     | -58                  | 4       |
| G8   | 384                  | 306     | O, H2, CIR 321                            | 321                  | 309     | -63                  | 3       |
| F7   | 530                  | 296     | A2, N4, W8                                | 468                  | 293     | -62                  | 3       |
| F7   | 530                  | 296     | A2, N4, M3                                | 461                  | 293     | -69                  | 3       |
| H9   | 348                  | 316     | E2.3, CIR, K1                             | 248                  | 327     | -100                 | 10      |
| H9   | 348                  | 316     | E2.3, G2, K1                              | 225                  | 327     | -123                 | 11      |



measurements are direct confirmation of changes in the velocity. We have derived only the average deceleration; we cannot say whether the rate of change is constant, increasing or decreasing. Knowledge of this would help to identify the physical cause of the deceleration, as would the development and application of quantitative models.

We thank James Foster, Dean Lindstrom, Doug MacAyeal, Richard Otto, Doug Roberts, Eric Roberts, John Scofield, Patricia Vornberger and the support staff of the NSF at Cray and Downstream B field camps who assisted in the collection of field data. A field team from the University of Wisconsin made the first determination of the position of station W8. We also thank C. R. Bentley for constructive criticism.

Received 20 January; accepted 28 June 1988.

- Hughes, T. J. *J. geophys. Res.* **78**, 7884–7910 (1973).
- Bindschadler, R. A., Stephenson, S. N., MacAyeal, D. R. & Shabtaie, S. *J. geophys. Res.* **92**, 8885–8894 (1987).
- Thomas, R. H., MacAyeal, D. R., Eilers, D. H. & Gaylord, D. R. *Antarct. Res. Ser.* **42**, 21–53 (1984).
- MacDonald, J. & Whillans, I. M. *Ann. Glaciol.* (in the press).
- Bindschadler, R. A. *et al. Ann. Glaciol.* (in the press).
- MacAyeal, D. R. in *Dynamics of the West Antarctic Ice Sheet* (eds van de Veen, C. J. & Oerlemans, J.) 141–160 (Reidel, Dordrecht, 1987).
- MacAyeal, D. R., Bindschadler, R. A., Shabtaie, S., Stephenson, S. N. & Bentley, C. R. *J. Glaciol.* **33**, 218–230 (1987).
- Shabtaie, S. & Bentley, C. R. *J. geophys. Res.* **92**, 1311–1336 (1987).
- Shabtaie, S., Bentley, C. R., Bindschadler, R. A. & MacAyeal, D. R. *Ann. Glaciol.* (in the press).
- Greischar, L. L. & Bentley, C. R. *Nature* **283**, 651–654 (1980).
- Jezek, K. C. *J. geophys. Res.* **89**, 409–416 (1984).

## Piracy as an alternative reproductive tactic for males

Eric P. van den Berghe

Department of Biological Sciences, University of California, Santa Barbara, California 9310, USA

Males of many animals exhibit multiple mating behaviours<sup>1–3</sup> but most of these alternatives merely enable smaller, competitively inferior males to obtain a few matings where they could get none in regular competition. Some alternatives may match the mating rate of dominant males<sup>4,5</sup>, but alternatives with higher payoffs than being a nesting or territorial male are virtually unknown. Here, I describe a new alternative, 'piracy', in which the largest males circumvent the costs of nest construction, guarding and potential nest failure by taking over successful nests of other males, spawning in these, and abandoning them to be guarded by the original owner. The overall payoff to a pirate male appears to be above that of the average nesting male and 2.5 to 10 times higher than that of the other documented alternatives in the population.

I studied the peacock wrasse *Symphodus tinca*, a common nearshore Mediterranean labrid fish, over three breeding seasons (April to late June in 1985–87) at the University of Liège marine biology station, STARESO, located in the Bay of Calvi, Corsica. Female *S. tinca* spawn daily, usually in the morning, and partition each day's eggs among many spawns of about 50 eggs each<sup>6</sup>. The eggs are adhesive and may be laid on algae-covered substrate either in or out of nests<sup>6</sup>. Several male mating behaviour patterns are known for *S. tinca*<sup>6–8</sup>, these include 'nesting' males, 'satellites', and 'interceptors'.

Nesting males generally exceed 20 cm (standard length, SL) and build simple algal nests on rocky substrates which they defend until their young hatch and disperse. One-third of nests (24/74) were 'unsuccessful' (never achieved high spawning rates and were abandoned prior to hatching). Successful nesting males averaged ( $\pm 1$  standard deviation)  $6.0 \pm 2.9$  spawns in 10 min

over the entire nest cycle, and sometimes got over a thousand spawns in the nest. Each cycle started with active nest construction, followed by high spawning rates ( $>20$  spawns in 10 min) which gradually decreased and stopped about halfway through the nest cycle. Parental care consisted of guarding the eggs until they hatched some 7 to 12 days after oviposition. Metabolic costs of nesting are substantial; some nesting males lost up to 20% of their body weight in a month, whereas 'interceptors' (below) lost no weight (van den Berghe, in preparation), and most nesting males spent less than half of the 50-day breeding season at nests (Table 1).

Satellites are smaller (10–18 cm SL), younger, individuals who gathered around the periphery of a nest to spawn by rushing into the nest whenever a female spawned. The average spawning rate of 10 satellites followed over the 1986 breeding season was  $0.5 \pm 0.4$  spawns in 10 min. 'Interceptors' (12–24 cm SL) frequented either rocky outcrops or nesting territories where many females passed. These males intercepted females who sometimes spawned away from nests where no parental care would be provided. Interceptors averaged  $1.9 \pm 0.7$  spawns in 10 min in 1986 ( $N = 17$  males).

Here, I describe a fourth alternative which I shall call 'piracy'. Pirates are the largest males ( $>22$  cm SL) and parasitize the reproductive effort of nesting males by fighting with nest owners and taking over actively spawning nests, only to abandon the nests to the care of the original owners when spawning ceases. Nest takeovers by pirates generally occurred after a nest had accumulated many eggs from several days of active spawning. A pirate would occupy each nest for only the last third of the nest's spawning period, during which he monopolized mating, and averaged  $5.0 \pm 0.8$  spawns in 10 min. This rate is comparable to that of nesting males (Table 1) and 2.5- to 10-fold greater than satellites or interceptors. The original nest owners continued to guard the eggs both while the pirate was in control and after he abandoned. Pirates themselves rarely defended against satellites, but often chased former owners (Table 1). Pirates fed at about twice the rate of nesting males, and did so primarily in nests (eating eggs), whereas nesting males fed almost exclusively outside nests (Table 1). Note, however, that

Table 1 Behaviour and characteristics of pirates, pirated nest owners, and normal nest owners.

|                                        | Pirate<br>( $N = 12$ ) | Pirated nest<br>owner ( $N = 12$ ) | Normal nest<br>owner ( $N = 25$ ) |
|----------------------------------------|------------------------|------------------------------------|-----------------------------------|
| Length (mm SL)                         | $245 \pm 9$            | $219 \pm 16$                       | $222 \pm 4$                       |
| Days at nest                           | $1.9 \pm 0.3$          | $19.2 \pm 9.3$                     | $23.5 \pm 10.0$                   |
| Days at nest during the spawning phase | $1.7 \pm 0.3$          | $5.7 \pm 1.5$                      | $5.4 \pm 2.1$                     |
| Nest construction                      | No                     | No                                 | Yes                               |
| Chased satellites                      | $0.3 \pm 0.2$          | $3.6 \pm 1.4$                      | $5.4 \pm 1.9$                     |
| Chased owner                           | $5.9 \pm 1.6$          | NA                                 | NA                                |
| Bites in nest                          | $4.7 \pm 0.9$          | 0                                  | $0.1 \pm 0.2$                     |
| Bites out of nest                      | $0.2 \pm 0.3$          | $2.7 \pm 0.9$                      | $2.3 \pm 1.5$                     |
| Spawning rate                          | $5.0 \pm 0.8$          | $0.1 \pm 0.1$                      | $6.0 \pm 3.9$                     |

All chasing, feeding (bites), and spawning values are mean rates per 10 min  $\pm$  one standard deviation. Pirated nest owner values encompass only the period during which the pirate was present; those for normal owners are averaged over the entire nest cycle and do not include pirated nests or those which were abandoned early.  $N$  is the number of males; only the mean for each male was used. Underlined values are not significantly different from another at the 0.05 level (ANOVA with Tukey procedure at  $\alpha = 0.05$ ). Observations were carried out on an individually marked wild population *in situ* using SCUBA. Marks consisted of Alcian Blue<sup>12</sup> injected into scale pockets. Each male was observed daily for 5 or 10 min per individual, sometimes over successive breeding seasons. Records were kept of location, feeding, nest defence, and spawning for each male. Spawns involving more than one male were divided equally among participant males<sup>7</sup>. Most pirates were not captured but their lengths were estimated in side-by-side comparisons with marked males. NA, not applicable.

pirate egg consumption tended to occur immediately after a nest takeover and at neighbouring nests, and rarely in nests where the pirate had spawned (data not shown).

Nest sites were not limiting: of the 19 known successful nest sites in the station harbour, no more than 11 were ever occupied simultaneously, and the average numbers occupied concurrently in a breeding season were  $7.7 \pm 0.5$ ,  $4.3 \pm 0.4$ , and  $4.2 \pm 0.4$ . There were, however, site preferences; the same three sites were among the first five to be occupied in all three seasons, and remained occupied for  $28 \pm 11$  days ( $N=9$ ), compared to  $18 \pm 10$  days (zeros excluded;  $N=47$ ) for all other successful sites ( $t=2.54$ ,  $P<0.01$  where  $t$  is derived in Student's  $t$ -test,  $P$  is probability of wrongly rejecting null hypothesis). However, these preferred sites were no more likely to be taken over by pirates. Only 11% (1/9) of the preferred sites were taken over, compared to 23% (11/47) of all other nests ( $\chi^2$  (degrees of freedom) = 0.06;  $P>0.5$ ).

Most males constructed nests within 100 m of their home ranges (van den Berghe, in preparation). However, pirates were rarely seen except when pirating a nest, and presumably came from outside the 500 m coastal zone (to 15 m deep) which was surveyed regularly. Nesting behaviour was only observed twice in known pirates. Both these pirates took over neighbouring nests after completing a cycle at a nest they had built and defended themselves. One of these built a second nest nearby after two days of piracy.

Nest takeovers were witnessed only three times; in one of these, the pirate visited eight nests in 25 min before taking one over after four minutes of fighting. The original owner resumed control on the following day when spawning stopped and the pirate abandoned. Of the seven nests rejected by the pirate, two were past the spawning phase, two were early in their cycles (one of these was pirated two days later by the same male), one already had a pirate, and the last two were not successful. The rapid nest assessment was possible because nests receiving active spawning often have associated females or satellites, and because eggs are visible.

A pirated nesting male is caught in a trap: if he abandons, the eggs in the nest will be preyed upon, and a new cycle takes several days to initiate. If many eggs in the nest are his, he should continue to care for them. However, if a pirate takes over a nest too early in the cycle, then the nesting male has relatively few prospects of getting enough of his own eggs to justify defence and should abandon. Consequently, takeovers generally occurred after the most active spawning period when nests contained sufficient numbers of eggs to make the nest worth defending by the original male. Only once did a pirate take over a nest with few eggs; here, the former owner abandoned first, followed by the pirate.

Although the spawning rate of pirates is comparable to that of nesting males, the net payoff is probably higher. This is because pirates do not expend a week or more of time and energy guarding between cycles, pirates can target already successful nests for takeover, (thereby avoiding the 32% of nests which are abandoned by nesting males), and pirates may be able to breed longer than most nesting males because they feed at twice the rate while they are breeding.

Given high payoffs and low risk to pirates, it would be surprising if this tactic were not common whenever it is possible to temporarily displace a successful nesting male. It is likely that the largest males of any nesting fish can displace somewhat smaller individuals. However, at least one requirement must be satisfied before other males will continue to care for the brood after the pirate leaves: either the cost of care must not increase greatly with offspring number, or breeding sites must be limiting, when it is less likely that a pirated nesting individual can abandon and be successful elsewhere. These preconditions exist in diverse freshwater and marine fishes where males who are removed from their nests are replaced by others who resume parental duties<sup>3,10-12</sup>. Other likely cases of piracy occur in

darters<sup>10-11</sup>, and in a related wrasse<sup>3</sup>. Further work may reveal piracy as an important alternative reproductive tactic for the largest males in other nesting fish species.

This work was funded by a Belgian American Educational Foundation fellowship to EvdB and an NSF grant to R. Warner. Valuable field assistance was provided by R. Warner and F. Wernerus. I also thank the staff of STARESO for logistic support, and J. Endler, T. Neal, S. Rothstein, E. Schultz, M. Taborsky and R. Warner for helpful suggestions on the manuscript.

Received 4 April; Accepted 7 July 1988.

1. Rubenstein, D. I. in *Limits to Action* (ed. Staddon, J. R.) 65-100 (Academic, New York, 1980).
2. Gross, M. R. in *Fish Reproduction: Strategies and Tactics* (eds Wootton, G. W. & Potts, R. J.) 55-75 (Academic, London, 1984).
3. Taborsky, M., Hudde, B. & Wirtz, P. *Behaviour* **102**, 82-117 (1987).
4. Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge University Press, 1982).
5. Gross, M. R. *Nature* **313**, 47-48 (1985).
6. Lejeune, P. *Cahier d'Ethologie Appliquée* **5**, 1-208 (1985).
7. Warner, R. R. & Lejeune, P. *Mar. Biol.* **87**, 89-99 (1985).
8. Michel, C. H., Lejeune, P. & Voss J. *Revue Française d'Aquariologie et Herpetologie* **14**, 1-80 (1987).
9. Constantz, G. D. *Proc. Acad. nat. Sci. Philad.* **131**, 131-138 (1979).
10. Constantz, G. D. *Envir. Biol. Fish.* **14**, 175-183 (1985).
11. Yasunobu, Y. & Ochi, H. *Anim. Behav.* **34**, 1769-1780 (1986).
12. Kelly, W. H. *Trans. Am. Fish. Soc.* **96**, 163-175 (1967).

## Interleukin-1 $\beta$ as a potent hyperalgesic agent antagonized by a tripeptide analogue

S. H. Ferreira\*, B. B. Lorenzetti\*, A. F. Bristow† & S. Poole†

\* Faculty of Medicine, 14.049 Ribeirão Preto, São Paulo, Brazil

† Department of Endocrinology, National Institute for Biological Standards and Control, Blanche Lane, South Mimms, Potters Bar, Herts EN6 3QG, UK

Interleukin-1 (IL-1) describes two inflammatory proteins<sup>1</sup>, IL-1 $\alpha$  and IL-1 $\beta$ , produced by activated macrophages and other cell types<sup>2</sup> and encoded by two genes. Their amino acid sequences have only 26% similarity<sup>1</sup>, but their biological activities are comparable, with a few exceptions<sup>3,4</sup>; indeed, both molecules appear to act at the same receptor<sup>5,6</sup>. As IL-1 releases prostaglandins<sup>7</sup> which sensitize nociceptors in man and in experimental animals<sup>8</sup>, we tested IL-1 $\alpha$  and IL-1 $\beta$  in rats for hyperalgesic (nociceptive) activity. Our results show that IL-1 $\beta$  given systemically is an extremely potent hyperalgesic agent with a probable peripheral site of action; IL-1 $\alpha$  is ~3,000 times less active than IL-1 $\beta$ . We have delineated the region of IL-1 $\beta$  mediating the hyperalgesic effect and developed an analgesic tripeptide analogue of IL-1 $\beta$  which antagonizes hyperalgesia evoked by IL-1 $\beta$  and by the inflammatory agent carrageenan.

Human recombinant (h.r.) IL-1 $\alpha$ , h.r. IL-1 $\beta$  and peptide analogues of IL-1 $\beta$  (Table 1) were tested for hyperalgesic/analgesic activities in a modified Randall-Sellito rat-paw pressure test<sup>9</sup>. Three of the peptides ( $\beta$ 121-134-Tyr,  $\beta$ 140-153 and Tyr- $\beta$ 190-202) had been synthesized originally for other purposes and two of them contained a 'non-native' tyrosine residue added at either the N- or C-terminus. Injection of IL-1 $\beta$  into one paw (intraplantar, i.p.l.) evoked a dose-dependent hyperalgesia in both paws, except for the smallest dose, which evoked ipsilateral hyperalgesia (Fig. 1a). This result was interpreted as indicating a systemic distribution of the injected IL-1 $\beta$ , an interpretation supported by the finding that bilateral hyperalgesia developed after intraperitoneal (i.p.) administration of IL-1 $\beta$  (Fig. 1c). Local pretreatment with indomethacin (an aspirin-like drug that



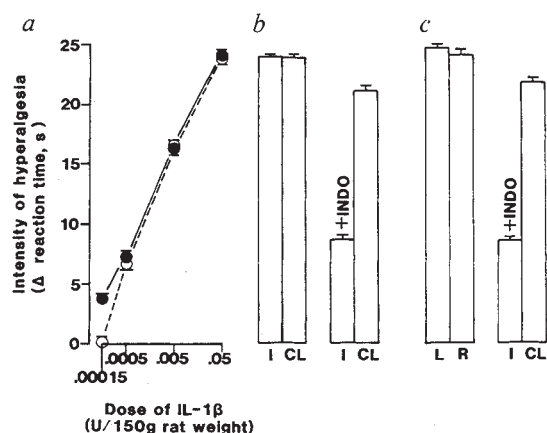
inhibits cyclo-oxygenase<sup>10</sup>) markedly attenuated the development of IL-1 $\beta$ -evoked hyperalgesia in the treated paws (Fig. 1b and c), suggesting that local release by IL-1 $\beta$  of cyclo-oxygenase products such as prostaglandins in the vicinity of nociceptors contributed to the hyperalgesia. (The small effect of indomethacin in contralateral paws may have been due to systemic distribution of a proportion of the indomethacin given i.p.).

Two of the peptide analogues of IL-1 $\beta$ , Tyr- $\beta$ 190-202 and  $\beta$ 187-204, also evoked bilateral hyperalgesia when given i.p. (although for clarity, results in only one paw are given in Figs 2 and 3). In contrast,  $\beta$ 121-134-Tyr,  $\beta$ 140-153,  $\beta$ 156-174 and  $\beta$ 225-243 possessed little or no hyperalgesic activity. In view of the hyperalgesic activity of  $\beta$ 187-204, two tripeptides contained in this sequence, Lys-Asp-Asp ( $\beta$ 190-192) and Lys-Pro-Thr ( $\beta$ 193-195), were tested and found to be hyperalgesic, but were much less potent than the parent peptide and did not attain the same maximum effect (Fig. 2). It is possible that one of the tripeptides, Lys-Pro-Thr, was acting as a partial agonist, because non-hyperalgesic doses of Lys-Pro-Thr (20 and 30  $\mu$ g per 150 g rat weight), given subcutaneously (s.c.) 30 min before IL-1 $\beta$  (0.05U, i.p.) markedly attenuated the hyperalgesic effect of IL-1 $\beta$ . In contrast, Lys-Asp-Asp given at a non-hyperalgesic dose (50  $\mu$ g per 150 g rat weight) before IL-1 $\beta$  (0.05U), had little effect on IL-1 $\beta$ -evoked hyperalgesia.

Modification of the structure of Lys-Pro-Thr to Lys-D-Pro-Thr gave a tripeptide that was almost devoid of intrinsic activity (Fig. 2), but which inhibited the bilateral hyperalgesic effects of IL-1 $\beta$  and  $\beta$ 187-204 in a dose-dependent manner. The ED<sub>50</sub> of Lys-D-Pro-Thr (given s.c.) for inhibiting hyperalgesia evoked by IL-1 $\beta$  (0.05U per 150 g rat weight, given i.p. 30 min later) was 85  $\mu$ g per 150 g rat weight. The inhibition was surmounted by increasing the dose of IL-1 $\beta$  (Fig. 3), which is consistent with a competitive antagonism. Also from Fig. 3, it can be seen that IL-1 $\beta$  was about 3,000 times more potent than IL-1 $\alpha$  as a hyperalgesic agent, and that IL-1 $\alpha$ , which lacks the sequence Lys-Pro-Thr, was only weakly antagonized by Lys-D-Pro-Thr.

The dose of Lys-D-Pro-Thr (200  $\mu$ g per 150 g) that antagonized hyperalgesia in response to IL-1 $\beta$  also antagonized the hyperalgesia evoked by injection into the paw of carrageenan, but not of PGE<sub>2</sub> (Fig. 3). The maximum analgesic effect of Lys-D-Pro-Thr was similar to that obtained with indomethacin<sup>9,11</sup>. The residual hyperalgesia after treatment with Lys-D-Pro-Thr was probably due to the sympathetic component present in carrageenan-evoked hyperalgesia<sup>12</sup>, as the combination of the adrenergic neuron-blocking agent guanethidine (4.5 mg per 150 g rat weight, given s.c. daily for 3 days) with Lys-D-Pro-Thr (200  $\mu$ g per 150 g given i.p. 30 min before the injection of carrageenan) abolished carrageenan-evoked hyperalgesia.

The possibility that the analgesic effect of Lys-D-Pro-Thr was mediated centrally may be excluded as the peptide did not antagonize PGE<sub>2</sub>-evoked hyperalgesia (Fig. 3), unlike centrally



**Fig. 1** Intensity of hyperalgesia 3 h after injection of IL-1 $\beta$  into one paw (a and b, injection volume 0.1 ml i.p.; c, injection volume 0.3 ml i.p.) and its attenuation by pretreatment with indomethacin (b and c, INDO injected i.p.). a, The dose-dependent hyperalgesic effect of IL-1 $\beta$  in the injected paws (I, filled circles and solid lines) and in contralateral paws (CL, open circles and dotted lines). b and c, Pretreatment with indomethacin (100  $\mu$ g in 0.1 ml given i.p.) markedly attenuated the hyperalgesic response to the injection 30 min later of IL-1 $\beta$  (0.05U) given i.p. (b) or i.p. (c) in the treated paws (I) (L, left paw and R, right paw). Hyperalgesia (nociception) was measured by a modification of the Randall-Sellito rat-paw pressure test<sup>9</sup>. The development of hyperalgesia was evaluated by the application of a constant pressure of 20 mm Hg to the hind paws of rats (Wistar strain, male, weight 135–170 g) which was discontinued when animals presented a characteristic freezing reaction (reaction time). The intensity of hyperalgesia was quantified as the variation of the reaction time ( $\Delta$  reaction time) obtained by subtracting the value measured 3 h after administration of the hyperalgesic agent from the pre-injection control reaction time (zero time). The experimenter was unaware of the group treatments.  $\Delta$  reaction times were maximal at intervals of 1, 2, 3 and 4 h after administration of IL-1, sub-maximal 6 h after administration and had returned to pre-injection values within 24 h. The vertical bars are standard errors of the mean (s.e.m.) of values obtained in groups of 5 rats. The IL-1 $\beta$ , and the IL-1 $\alpha$  used in other experiments, were human recombinant proteins purchased from Genzyme Ltd and calibrated against the Interim Reference Reagents (temporary biological standards from NIBSC) for IL-1 $\alpha$  and IL-1 $\beta$ , which have assigned potencies of 1 unit per 10 pg.

acting analgesics such as morphine<sup>11</sup>. Also, Lys-D-Pro-Thr (32 mg kg<sup>-1</sup> given s.c.) was not active in a 'hot-plate test' (55 °C) in mice, in contrast to morphine (3 mg kg<sup>-1</sup>, s.c.). The possibility that Lys-D-Pro-Thr was interfering with the synthesis and release of prostaglandins by inhibiting cyclo-oxygenase was excluded by demonstrating that Lys-D-Pro-Thr (200  $\mu$ g ml<sup>-1</sup>), in contrast to the cyclo-oxygenase inhibitor indomethacin<sup>10</sup> (2  $\mu$ g ml<sup>-1</sup>),

**Table 1** Amino acid sequences of analogues of IL-1 $\beta$  tested for hyperalgesic/analgesic activity and their positions within the IL-1 $\beta$  sequence

| Residue numbers       | Peptide sequence                          |   |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |
|-----------------------|-------------------------------------------|---|---|---|---|---|-------|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ 113-271      | IL-1 $\alpha$ mature protein <sup>1</sup> |   |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |
| $\beta$ 117-269       | IL-1 $\beta$ mature protein <sup>1</sup>  |   |   |   |   |   |       |   |   |   |   |   |   |   |   |   |   |   |   |
| $\beta$ 121-134 (TYR) | S                                         | L | N | C | T | L | R     | D | S | Q | Q | K | S | L | Y |   |   |   |   |
| $\beta$ 140-153       | Y                                         | E | L | K | A | L | H     | L | Q | G | Q | D | M | E |   |   |   |   |   |
| $\beta$ 156-174       | V                                         | V | F | S | M | S | F     | V | Q | G | E | E | S | N | D | K | I | P | V |
| $\beta$ 187-204       | C                                         | V | L | K | D | D | K     | P | T | L | Q | L | E | S | V | D | P | K |   |
| (TYR) $\beta$ 190-202 |                                           |   | Y | K | D | D | K     | P | T | L | Q | L | E | S | V | D |   |   |   |
| $\beta$ 190-192       |                                           |   |   | K | D | D |       |   |   |   |   |   |   |   |   |   |   |   |   |
| $\beta$ 193-195       |                                           |   |   |   |   |   | K     | P | T |   |   |   |   |   |   |   |   |   |   |
|                       |                                           |   |   |   |   |   | K (D) | P | T |   |   |   |   |   |   |   |   |   |   |
| $\beta$ 225-243       | K                                         | L | E | F | E | S | A     | Q | F | P | N | W | Y | I | S | T | S | Q | A |

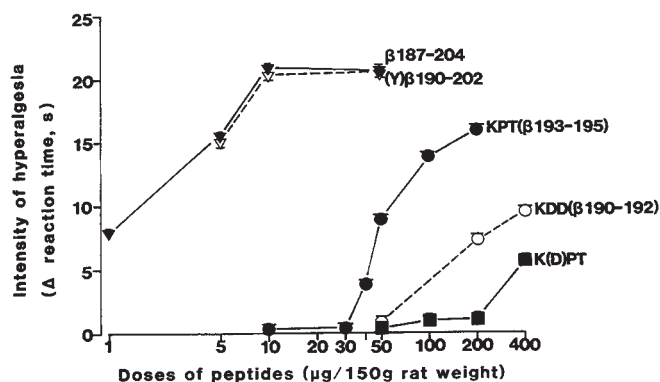


Fig. 2 Hyperalgesic responses of rats to peptide analogues of IL-1 $\beta$ . KDD represents Lys-Asp-Asp; KPT, Lys-Pro-Thr and K(D)PT, Lys-D-Pro-Thr. The effects were measured 3 h after i.p. injection. Each symbol represents the mean response of 5 animals per treatment group, vertical bars, the s.e.m. The peptides were synthesized by Cambridge Research Biochemicals (Cambridge, England) and Peninsula Laboratories (Merseyside, England) and purified by preparative HPLC. Identity and purity were established by fast atom bombardment mass spectrometry, amino acid analysis and analytical reverse-phase HPLC.

did not diminish basal release of PGE<sub>2</sub> by human blood mononuclear cells ( $5 \times 10^6$  per ml) or release stimulated by incubation for 4 h with IL-1 $\alpha$  (10–50 U per ml), IL-1 $\beta$  (10–50 U per ml) and endotoxin (25–100 pg per ml).

Lys-D-Pro-Thr and the other peptides given in Table 1 were tested for agonist/antagonist activity in an *in vitro* EL-4 thymoma conversion assay for IL-1 (ref. 13) and in the rabbit pyrogen test<sup>14</sup>. In these systems neither Lys-D-Pro-Thr nor any of the other eight peptides showed agonist or antagonist activity when tested at doses of up to  $200 \mu\text{g ml}^{-1}$  and  $200 \mu\text{g kg}^{-1}$  (i.v.) respectively. In the rabbit pyrogen test, the peptides  $\beta$ 121–134-Tyr,  $\beta$ 140–153 and Tyr- $\beta$ 190–202 were also inactive when given at  $30 \mu\text{g kg}^{-1}$  into the cerebral ventricles, and Lys-D-Pro-Thr did not have a consistent inhibitory effect on IL-1 $\beta$ -evoked fever when given at doses of 1 and  $2 \text{ mg kg}^{-1}$  (i.v.), either with the IL-1 $\beta$ , or 30 min or 60 min before the IL-1 $\beta$  ( $2.5 \times 10^4$ – $2 \times 10^5 \text{ U kg}^{-1}$  given i.v.).

The hyperalgesia evoked in rats by small doses of h.r. IL-1 $\beta$  ( $\leq 0.3 \text{ U kg}^{-1}$  i.p.l. or i.p.) and h.r. IL-1 $\alpha$  ( $\leq 300 \text{ U kg}^{-1}$  i.p.) contrasts with a recent report that much larger doses of h.r. IL-1 $\alpha$  evoked analgesia in mice by a central action<sup>15</sup> ( $\text{ED}_{50} = 5 \times 10^3 \text{ U kg}^{-1}$ , intracisternal;  $2 \times 10^5 \text{ U kg}^{-1}$ , i.v.;  $2 \times 10^6 \text{ U kg}^{-1}$ , i.p.). In rats  $<0.1\%$  of doses of up to  $10^4 \text{ U kg}^{-1}$  <sup>125</sup>I-labelled IL-1 given i.v. was found in brain at 10 min, 30 min and 1 h after injection (S.P., T. A., Bird and J. Saklatvala, manuscript in preparation), so it is unlikely that any significant quantity of the small doses of IL-1 injected i.p.l. or i.p. reached the brain in the present study. Also, IL-1 when injected i.v. into cats was not detectable subsequently in cerebrospinal fluid<sup>16</sup>.

Lys-D-Pro-Thr may be considered as the prototype for a new class of analgesic drugs; namely drugs that antagonize IL-1 $\beta$ . In contrast to steroidal and non-steroidal analgesics and anti-

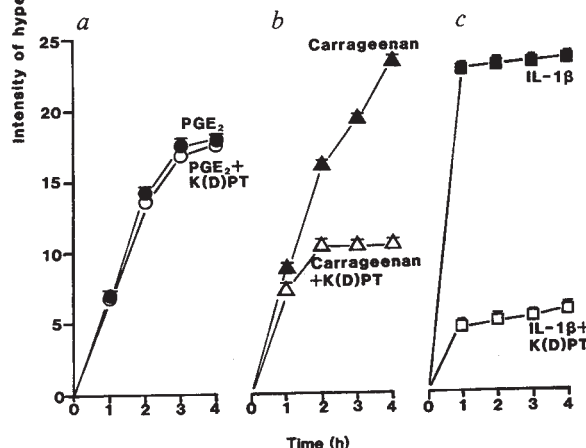
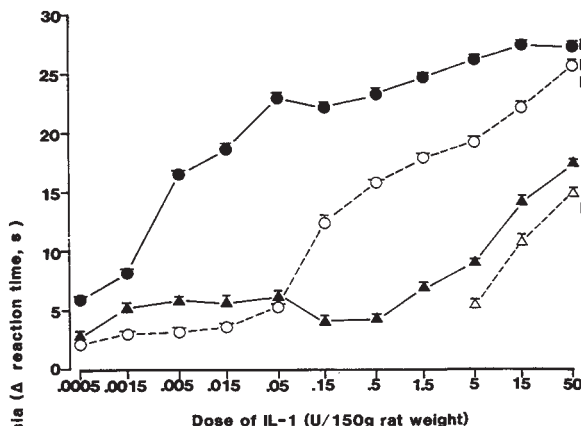


Fig. 3 Upper panel: the hyperalgesic responses to IL-1 $\alpha$  and IL-1 $\beta$  and the effect of K(D)PT (Lys-D-Pro-Thr) on those responses. Lys-D-Pro-Thr ( $200 \mu\text{g}$  per  $150 \text{ g}$  rat weight) was given s.c., 30 min before the i.p. injection of IL-1. Intensity of hyperalgesia was measured 3 h after injection of IL-1. Lower panel: effect of K(D)PT on the temporal development of hyperalgesic responses to a, PGE<sub>2</sub> ( $100 \text{ ng}$  per paw); b, carrageenan ( $100 \mu\text{g}$  per paw) and c, IL-1 $\beta$  ( $0.05 \text{ U}$  per paw). Lys-D-Pro-Thr ( $200 \mu\text{g}$  per  $150 \text{ g}$ ) was injected i.p., 30 min before the i.p.l. injections. Each symbol represents the mean response of 5 animals per treatment group, vertical bars represent s.e.m.

inflammatories<sup>7,17</sup>, Lys-D-Pro-Thr was not an effective antipyretic against IL-1 $\beta$  and did not antagonize the oedema evoked by carrageenan (measured plethysmographically<sup>18</sup> 4 h after the challenge). Nevertheless, Lys-D-Pro-Thr (given s.c., i.p. or i.p.l.) was a very effective antagonist of IL-1 $\beta$  in the modified Randall-Sellitto test<sup>9</sup>. Also, the lack of effect of Lys-D-Pro-Thr on PGE<sub>2</sub> production suggests that Lys-D-Pro-Thr and its congeners will not cause the gastric lesions that limit the usefulness in man of non-steroidal analgesics.

The authors thank I. R. Santos for technical assistance.

Received 6 April; accepted 18 July 1988.

- March, C. J. *et al.* *Nature* **315**, 641–646 (1985).
- Oppenheim, J. H., Kovacs, E. J., Matsushima, K. & Durum, S. K. *Immunol. Today* **7**, 45–56 (1986).
- Uehara, A., Gottschall, P. E., Dahl, R. R. & Arimura, A. *Biochem. biophys. Res. Commun.* **146**, 1286–1290 (1987).
- Lumpkin, M. D. *Science* **238**, 452–454 (1987).
- Bird, T. A. & Saklatvala, J. *Nature* **324**, 263–266 (1986).
- Dower, S. K. *et al.* *Nature* **324**, 266–268 (1986).
- Bernheim, H. A., Gilbert, T. M. & Stitt, J. T. *J. Physiol.* **301**, 69–78 (1980).
- Ferreira, S. H. *Nature new Biol.* **240**, 200–203 (1972).

- Ferreira, S. H., Lorenzetti, B. B. & Correa, F. M. A. *Eur. J. Pharmac.* **53**, 39–48 (1978).
- Vane, J. R. *Nature* **231**, 232–235 (1971).
- Lorenzetti, B. B. & Ferreira, S. H. *Eur. J. Pharmac.* **114**, 375–381 (1985).
- Nakamura, M. & Ferreira, S. H. *Eur. J. Pharmac.* **135**, 145–153 (1987).
- Gearing, A. J. H., Bird, C. R., Bristow, A. F., Poole, S. & Thorpe, R. *J. immun. Meth.* **99**, 7–11 (1987).
- Test for Pyrogens British Pharmacopoeia II*. Appendix XIV K A183–A184, (HMSO, London, ISBN 0 11 320837 5, 1988).
- Nakamura, H., Nakanishi, K., Kita, A. & Kadokawa, T. *Eur. J. Pharmac.* **149**, 49–54 (1988).
- Cocconi, F., Lees, J. & Dinarello, C. A. *Brain Res.* **446**, 245–250 (1988).
- Willies, G. H. & Woolf, C. J. *J. Physiol.* **300**, 1–6 (1980).
- Ferreira, S. H. *J. Pharm. Pharmac.* **31**, 648 (1979).



# Neural adhesion molecule L1 as a member of the immunoglobulin superfamily with binding domains similar to fibronectin

Marion Moos\*, Roland Tacke\*, Herta Scherer\*,  
David Teplow†, Klaus Früh\* & Melitta Schachner\*

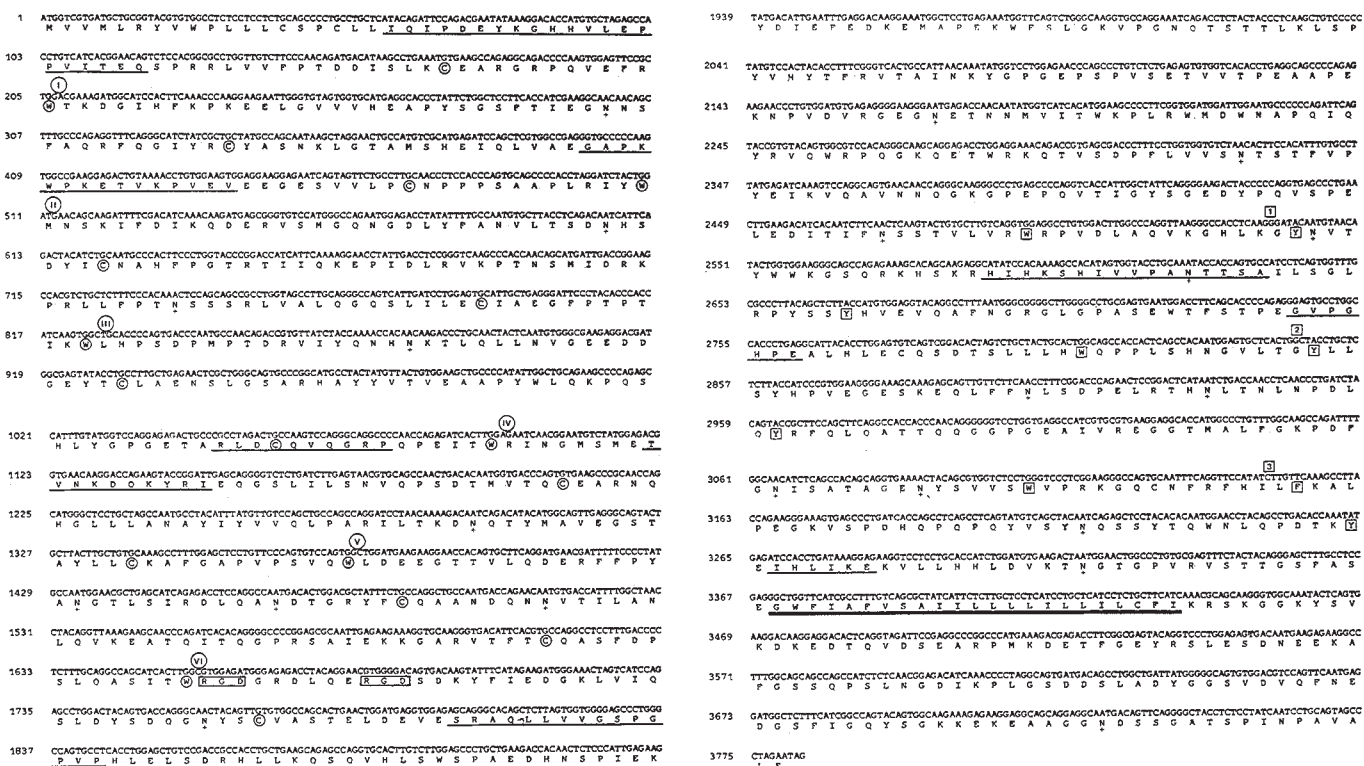
\* Department of Neurobiology, University of Heidelberg,  
Im Neuenheimer Feld 364, 6900 Heidelberg, FRG

† Division of Biology 147-75, California Institute of Technology,  
Pasadena, California 91125, USA

Diverse glycoproteins of cell surfaces and extracellular matrices operationally termed 'adhesion molecules' are important in the specification of cell interactions during development, maintenance and regeneration of the nervous system<sup>1,2</sup>. These adhesion molecules have distinct functions involving different cells at different developmental stages, but may cooperate when expressed together<sup>3-6</sup>. Families of adhesion molecules which share common carbohydrate domains do exist, despite the structural and functional diversity of these glycoproteins<sup>7,8,9</sup>. These include the Ca<sup>2+</sup>-independent neural adhesion molecules: N-CAM, myelin associ-

ated glycoprotein (MAG)<sup>10</sup> and L1. L1 is involved in neuron-neuron adhesion<sup>6</sup>, neurite fasciculation<sup>11</sup>, outgrowth of neurites<sup>12</sup>, cerebellar granule cell migration<sup>13</sup>, neurite outgrowth on Schwann cells<sup>14</sup> and interactions among epithelial cells of intestinal crypts<sup>15</sup>. We show here that in addition to sharing carbohydrate epitopes with N-CAM and MAG, L1 is also a member of the immunoglobulin superfamily. It contains six C2 domains and also shares three type III domains with the extracellular matrix adhesion molecule fibronectin.

We previously described the isolation and partial characterization of complementary DNA clones of the neural cell adhesion molecule L1<sup>16</sup>. Further clones were selected and analysed by restriction mapping and hybridization with oligonucleotide probes derived from amino acid sequences of the amino-terminal end of two proteolytic fragments of L1, of relative molecular mass (*M<sub>r</sub>*) 80,000 (80K) and 140,000 (140K)<sup>17,18</sup>. Several clones were obtained that contained the complete nucleotide sequence of L1 of about 5,100 nucleotides, with an open reading frame of 3,783 residues (Fig. 1). Twenty-one amino acids deduced from the cDNA sequence, which were identical to those determined from the amino terminus of the protein, begin 19 amino acids downstream from the initiation codon. The protein contains 1,241 amino acids and has *M<sub>r</sub>* of 138,531, which is in good agreement with the biochemically determined value of approximately 150,000 for the deglycosylated glycoprotein<sup>17</sup>. In



**Fig. 1** Nucleotide sequence of the coding region of L1 with its predicted 1,260 amino acid sequence. The untranslated nucleotides (27 at the 5' and about 1,300 at the 3' ends) are not shown. The initiation codon is followed by a peptide consisting of predominantly hydrophobic amino acids, which is probably a signal peptide. The underlined sequences correspond to those determined by amino acid sequencing of the amino termini of the 140K and 80K proteolytic fragments of L1, as well as from sequencing peptides obtained after V8 protease treatment. A segment with 23 hydrophobic amino acid residues, presumably a transmembrane domain, is underlined by a bar. Potential sites of asparagine-linked glycosylation (Asn-X-Ser/Thr) are marked with (+). The two RGD sequences (putative cell attachment sites) are boxed. The immunoglobulin domains are numbered from I to VI and the characteristic cysteine and tryptophan are indicated by circles. The regions of homology with the fibronectin repeating unit type III are numbered 1-3 and the characteristic typtophan and tyrosine (or phenylalanine) residues are indicated by squares.

**Methods.** The first clone was isolated from a  $\lambda$ gt11 library of 8-day-old mouse brain cDNA constructed using oligo(dT) priming as described<sup>16</sup>. Further clones were obtained from a  $\lambda$ gt10 library constructed with the same cDNA as the  $\lambda$ gt11 library by hybridization with the first probe<sup>16</sup> and with additional oligonucleotide probes derived from peptide sequences. The nucleotide sequence was determined by the dideoxy-method<sup>20</sup> after subcloning the appropriate restriction fragments in M13-derived vectors<sup>21</sup>. Analyses of DNA and protein sequence data were carried out using the PCgene programme (Genofit).

|       |     |     |               |   |                       |       |         |           |           |               |               |               |               |
|-------|-----|-----|---------------|---|-----------------------|-------|---------|-----------|-----------|---------------|---------------|---------------|---------------|
| L1    | I   | 50  | T D D I S L K | C | E A R G R P Q V E F R | -     | W       | T K D G I | 31aa      | R F Q G I Y R | C             | Y A S N K L G |               |
|       | II  | 150 | G E S V V L P | C | N P P P S A A P       | -     | L R I Y | W         | M N S K I | 25aa          | L N H S D Y I | C             | N A H F P G T |
|       | III | 256 | G Q S L I L E | C | I A E G F P T P       | -     | T I K   | W         | L H P S D | 23aa          | E D D G E Y T | C             | L A E N S L G |
|       | IV  | 346 | G E T A R L D | C | Q V Q G R P Q P E     | -     | I T     | W         | R I N G M | 25aa          | S D T M V T Q | C             | E A R N Q H G |
|       | V   | 440 | G S T A Y L L | C | K A F G A P V P S     | Q V   | -       | W         | L D E E G | 24aa          | N D T G R Y F | C             | Q A A N D Q N |
|       | VI  | 531 | G A R V T F T | C | Q A S F D P S L Q A   | S I T | W       | R G D G R | 24aa      | S D Q G N Y S | C             | V A S T E L D |               |
|       |     |     |               |   |                       |       |         |           |           |               |               |               |               |
| N-CAM | III | 228 | G Q S V T L V | C | D A D G F P E P T M S | -     | W       | T K D G E | 28aa      | N D E A E Y V | C             | I A E N K A G |               |
|       | IV  | 323 | E E Q V T L T | C | E A S G D P I P S     | -     | I T     | W         | R T S T R | 31aa          | R D A G E Y M | C             | T A S N T I G |
|       |     |     |               |   |                       |       |         |           |           |               |               |               |               |
| MAG   | III | 254 | G S H V S L L | C | G A D S N P P P       | -     | I I T   | W         | M R D G M | 19aa          | A E D G I Y A | C             | L A E N A Y G |
|       | IV  | 340 | G E T V S I L | C | S T Q S N P D P       | -     | I L T   | H         | F K E K Q | 20aa          | E D D G E Y W | C             | V A E N Q Y G |

**Fig. 2** Alignment (aa, amino acid) of L1 sequences around the conserved cysteines and tryptophans involved in the immunoglobulin fold and comparison with two of the corresponding regions of N-CAM and MAG. Boxed amino acids correspond to residues shared by two or more of the L1 domains or by one or more of the L1, MAG or N-CAM sequences.

addition to the signal peptide sequence between the first ATG codon and the amino terminus, the deduced 1,260 amino acids include a stretch of 23 hydrophobic residues characteristic of a membrane-spanning region (underlined by a black bar in Fig. 1) and 22 sites for asparagine-linked glycosylation (Asn-X-Ser/Thr) which are distributed over the molecule (marked with (+) in Fig. 1), with only one site located intracellularly.

A striking structural feature of L1 is the presence of six immunoglobulin-like domains (I–VI in Fig. 1) each containing approximately 50 amino acids between the two cysteines. The immunoglobulin domains are located towards the amino terminus with the first predicted domain starting about 15 amino acids from the amino terminus. The immunoglobulin-like domains belong to the C2 subgroup<sup>19</sup> and comprise approximately half of the coding sequence. Figure 2 shows the homology among the immunoglobulin domains of L1, as well as their homology with domains III and IV of N-CAM and MAG. Two Arg-Gly-Asp (RGD) sequences (boxed in Fig. 1) known to be involved in cell-binding events between extracellular matrix glycoproteins and cell surface receptors<sup>20</sup> are also found close together (6 amino acids apart) in domain VI.

A region in the third immunoglobulin domain (residues 261–275) is similar to a sequence in the tyrosine kinase transforming protein of feline sarcoma virus<sup>21</sup>, receptor for platelet-derived growth factor<sup>22</sup> and macrophage-colony-stimulating factor<sup>23</sup>. This is more pronounced with L1, N-CAM and MAG than with other members of the immunoglobulin superfamily.

There are three domains (1–3 in Fig. 1) which are similar to the 'repeating unit type III' of fibronectin, an extracellular matrix glycoprotein and adhesion molecule<sup>24,25</sup>. The three domains are close to the transmembrane region with approximately 30 amino acids between them. The third unit is 30 amino acids away from the hydrophobic transmembrane region. These homology regions are aligned in Fig. 3 with three of the type III repeating units of rat fibronectin whose characteristic conserved amino

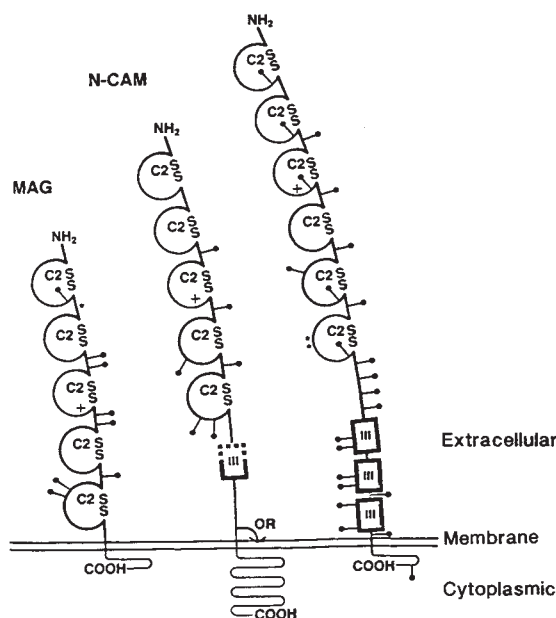
acids are tryptophan and tyrosine (or the conservative substitute phenylalanine). The sequence between the immunoglobulin domains and the membrane which also comprises the type III repeating units is rich in proline.

A schematic representation of the structural features of L1, N-CAM and MAG, is shown in Fig. 4. The sequence of L1 has two features of interest in the context of its functional properties as an adhesion molecule. Like N-CAM and MAG<sup>26–30</sup>, L1 is a member of the immunoglobulin superfamily. The domains of these three Ca<sup>2+</sup>-independent molecules belong to the C2 type which is characteristic of the lymphocyte adhesion molecules LFA3 and CD2, the CD3 chains of the T-cell receptor complex, the Fc (crystallizable fragment)-receptor, platelet-derived growth factor receptor<sup>19</sup> and I-CAM<sup>31</sup>. (Three Ca<sup>2+</sup>-dependent adhesion molecules whose primary structure is now available, E-cadherin<sup>32</sup> (uvomorolin, L-CAM), P-cadherin<sup>33</sup> and N-cadherin<sup>34</sup> do not belong to the immunoglobulin superfamily.) The functional significance of the immunoglobulin domains is unknown, but members of the immunoglobulin superfamily interact with each other. MHC class I antigens and CD1 interact with  $\beta_2$  microglobulin. Furthermore, the poly-immunoglobulin and the Fc-receptors interact with immunoglobulins and LFA3 interacts with CD2<sup>19</sup>. It is possible that these interactions occur through the immunoglobulin domains. The CH2 domains of the immunoglobulin heavy chains have a conserved N-glycosylation site with two oligosaccharide units in direct contact with each other, thus bridging the two protein domains<sup>35,36</sup>. For the two oligosaccharides to interact with each other, they must be asymmetric<sup>37</sup>. It may be that the carbohydrate moieties of adhesion molecules mediate cell binding<sup>6,9,38</sup> by similar arrangements of carbohydrates within immunoglobulin domains. The cell-binding domain of N-CAM has indeed been localized to the amino-terminal region containing the immunoglobulin domains<sup>39</sup> and the L2/HNK-1 carbohydrate moiety<sup>38</sup>. Furthermore, the unusual polysialic acid characteristic of the embryonic

|             |        |      |           |   |                   |             |   |             |      |                 |   |             |
|-------------|--------|------|-----------|---|-------------------|-------------|---|-------------|------|-----------------|---|-------------|
| L1          | III- 1 | 827  | T V L V R | W | R P V D L A Q V K | G H L K G   | Y | N V T Y W W | 28aa | S G L R P Y S S | Y | H V E V Q A |
|             | III- 2 | 932  | S L L L H | W | Q P P - L S H N   | - G V L T G | Y | L L S Y H P | 23aa | T N L N P D L Q | Y | R F Q L Q A |
|             | III- 3 | 1032 | S V V - S | W | V P R K G Q C N   | F R H I L - | F | K A L P E G | 22aa | W N L Q P D T K | Y | E I H L I K |
|             |        |      |           |   |                   |             |   |             |      |                 |   |             |
| Fibronectin | III- 5 | 981  | T V L V R | W | T P P - - - -     | R A R I A G | Y | R L T V G L | 20aa | R N L Q P G S E | Y | T V T L M A |
|             | III-10 | 1432 | S L L I S | W | E P P A V S - -   | - V R Y - - | Y | R I T Y G E | 21aa | N N I K P G A D | Y | T I T L Y A |
|             | III-14 | 1887 | S L L V S | W | Q A P - - - -     | R A R I T G | Y | I I K Y E K | 21aa | T G L E P G T E | Y | T I Y V I A |

**Fig. 3** Alignment of L1 sequences around the characteristic tryptophans, tyrosines (or phenylalanines) of the fibronectin repeating unit type III and comparison with three such repeating units of fibronectin. Boxed amino acids correspond to residues shared by two or more of the L1 or fibronectin domains or are common between L1 and fibronectin.





**Fig. 4** Topological comparison of the neural cell adhesion molecules L1, N-CAM and MAG. The immunoglobulin domains of type C2 are indicated with their characteristic disulphide bonds. The regions of L1 homologous to the fibronectin repeating unit type III and the truncated type III homology region in N-CAM are indicated with boxes. Possible glycosylation sites are indicated with the symbol (●). The RGD sequences found in L1 and MAG are marked with (\*). The homology between sequences of the tyrosine-kinase transforming protein of feline sarcoma virus and a part of the third immunoglobulin domain of L1, N-CAM and MAG is indicated by (+).

form of N-CAM is found in immunoglobulin domains IV and V.

The second important feature of L1 is its homology with fibronectin. The homologous domains form the binding domains for, at least, heparin and cells<sup>40</sup>. A truncated type III homologous domain was observed between the immunoglobulin and transmembrane domains of N-CAM<sup>26</sup>. As MAG does not contain type III domains, but interacts with heparin<sup>41,42</sup>, and the heparin-binding site for N-CAM has been localized to the immunoglobulin domains<sup>43</sup>, the type III domains of L1 may have additional functions. The presence of type III homologous domains is interesting in the light of previous findings that L1 (which is immunochemically identical to NILE<sup>44</sup> and Ng-CAM<sup>45</sup>) is found in the supernatants of cultured peripheral, but not central, nervous system neurons<sup>46,47</sup> and is detectable in basement membranes and on collagen fibres in developing, adult and regenerating sciatic nerve<sup>48-49</sup>. It is possible that these domains engage in interactions with extracellular matrix, perhaps not only in secreted but also in membrane-bound molecules. Thus distinctions between cell adhesion and substrate adhesion molecules could become blurred.

We are indebted to Peter Seeburg for help with DNA sequencing, W. Dreyer for encouragement and support, Patricia Kahn for helpful discussions, Bettina Crusius for technical assistance, Christo Goridis for comments on the manuscript and Bundesministerium für Forschung und Technologie for support.

Received 21 March; accepted 12 July 1988.

1. Edelman, G. M. & Thiery, J.-P. *The Cell in Contact* (Wiley, New York, 1985).
2. Edelman, G. M. & Thiery, J.-P. *The Cell in Contact* (Wiley, New York, in the press).
3. Thor, G., Pollerberg, E. G. & Schachner, M. *Neurosci. Lett.* **66**, 121-126 (1986).
4. Rathjen, F. G. & Rutishauser, U. *EMBO J.* **3**, 461-465 (1984).
5. Faissner, A., Kruse, J., Goridis, C., Bock, E. & Schachner, M. *EMBO J.* **3**, 733-737 (1984).
6. Keilhauer, G., Faissner, A. & Schachner, M. *Nature* **316**, 728-730 (1985).
7. Kruse, J., Keilhauer, G., Faissner, A., Timpl, R. & Schachner, M. *Nature* **316**, 146-148 (1984).
8. Kücherer, A., Faissner, A. & Schachner, M. *J. Cell Biol.* **104**, 1597-1602 (1987).
9. Künemund, V. et al. *J. Cell Biol.* **106**, 213-223 (1988).
10. Poltorak, M. et al. *J. Cell Biol.* **105**, 1893-1899 (1987).
11. Fischer, G., Künemund, V. & Schachner, M. *J. Neurosci.* **6**, 605-612 (1986).
12. Chang, S., Rathjen, F. G. & Raper, J. A. *J. Cell Biol.* **104**, 355-362 (1987).
13. Lindner, J., Rathjen, F. G. & Schachner, M. *Nature* **305**, 427-430 (1983).
14. Seilheimer, B. & Schachner, M. *J. Cell Biol.* (in the press).
15. Thor, G., Probstmeier, R. & Schachner, M. *EMBO J.* **6**, 2581-2586 (1987).
16. Tacke, R. et al. *Neurosci. Lett.* **82**, 89-94 (1987).
17. Faissner, A. et al. *EMBO J.* **4**, 3105-3113 (1985).
18. Sadoul, K., Sadoul, R., Faissner, A. & Schachner, M. *J. Neurochem.* **50**, 510-521 (1988).
19. Williams, A. F. *Immunol. Today* **8**, 298-303 (1987).
20. Ruoslahti, E. & Pierschbacher, M. D. *Cell* **44**, 517-518 (1986).
21. Hampe, E., Gobet, M. & Galibert, F. *Proc. natn. Acad. Sci. U.S.A.* **81**, 85-89 (1984).
22. Yarden, Y. et al. *Nature* **323**, 226-232 (1986).
23. Sherr, C. J. et al. *Cell* **41**, 665-676 (1985).
24. Hynes, R. O. *Scient. Am.* **254**, 42-51 (1986).
25. Hynes, R. O. *A. Rev. Cell Biol.* **1**, 67-90 (1985).
26. Hemperly, J. J., Murray, B. A., Edelman, M. G. & Cunningham, B. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3037-3041 (1986).
27. Arquint, M. et al. *Proc. natn. Acad. Sci. U.S.A.* **84**, 600-604 (1987).
28. Barthels, D. et al. *EMBO J.* **6**, 907-910 (1987).
29. Lai, C. et al. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4337-4341 (1987).
30. Salzer, J. L., Holmes, W. P. & Colman, D. R. *J. Cell Biol.* **104**, 957-965 (1987).
31. Simmons, D., Makgoba, M. W. & Seed, B. *Nature* **331**, 624-627 (1988).
32. Nagafuchi, A., Shirayoshi, Y., Okazaki, K., Yasuda, K. & Takeichi, M. *Nature* **329**, 341-343 (1987).
33. Nose, A., Nagafuchi, A. & Takeichi, M. *EMBO J.* **6**, 3655-3661 (1987).
34. Hatta, K., Nose, A., Nagafuchi, A. & Takeichi, M. *J. Cell Biol.* **106**, 873-881 (1988).
35. Parekh, R. B., Dwek, R. A., Sutton, B. J., Fernandes, D. L. & Leung, A. *Nature* **316**, 452-457 (1985).
36. Sutton, B. J. & Phillips, D. C. *Biochem. Soc. Trans.* **11**, 130-132 (1982).
37. Rademacher, T. W., Homans, S. W., Parekh, R. B. & Dwek, R. A. *Biochem. Soc. Symp.* **51**, 131-148 (1986).
38. Cole, G. J. & Schachner, M. *Neurosci. Lett.* **78**, 227-232 (1987).
39. Cunningham, B. A., Hoffman, S., Rutishauser, U., Hemperly, J. J. & Edelman, G. M. *Proc. natn. Acad. Sci.* **80**, 3116-3120 (1983).
40. Schwarzbauer, J. E., Patel, R. S., Fonda, D. & Hynes, R. O. *EMBO J.* **6**, 2573-2580 (1987).
41. Fahrig, T., Landa, C., Pesheva, P., Kühn, K. & Schachner, M. *EMBO J.* **6**, 2875-2883 (1987).
42. Cole, G. J., Loewy, A. & Glaser, L. *Nature* **320**, 445-447 (1986).
43. Cole, G. J. & Glaser, L. *J. Cell Biol.* **102**, 402-412 (1986).
44. Bock, E., Richter-Landsberg, C., Faissner, A. & Schachner, M. *EMBO J.* **4**, 2765-2768 (1985).
45. Friedlander, D. R., Grumet, M. & Edelman, G. M. *J. Cell Biol.* **102**, 413-419 (1986).
46. Sweadner, K. J. *J. Neurosci.* **3**, 2504-2517 (1983).
47. Sweadner, K. J. *J. Neurosci.* **3**, 2518-2524 (1983).
48. Martini, R. & Schachner, M. *J. Cell Biol.* **103**, 2439-2448 (1986).
49. Martini, R. & Schachner, M. *J. Cell Biol.* **106**, 1735-1746 (1988).
50. Sanger, F., Nicklin, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
51. Messing, J. *Meth. Enzym.* **101**, 20-78 (1983).

## Calcium channels in thrombin-activated human platelet membrane

A. Zschauer, C. van Breemen\*, F. R. Bühler & M. T. Nelson†

Departments of Research and Medicine, University Hospital, Basel, Switzerland

\* Pharmacology Department, University of Miami Medical School, Florida 33101, USA

† University of Vermont College of Medicine, Department of Pharmacology, Burlington, Vermont 05403, USA

Platelet-activating factor, 5-hydroxytryptamine, thromboxane A<sub>2</sub>, adenosine diphosphate and thrombin are known to activate platelets by stimulating calcium entry<sup>1,2</sup>, but the nature of the entry pathways is unknown<sup>3</sup>. We present the identification of single divalent cation channels from thrombin-activated human platelets. Membrane vesicles from unstimulated and thrombin-stimulated human platelets were incorporated in planar bilayers and unitary currents through single channels were measured. Divalent cation selective channels could only be demonstrated in thrombin-stimulated preparations. These channels share a number of properties in common with voltage-dependent calcium channels—a high degree of selectivity for divalent cations, a single channel conductance of about 10 pS (in 150 mM Ba<sup>2+</sup>) and sensitivity to blockade by inorganic calcium channel blockers such as Ni<sup>2+</sup>. In other respects, these channels are different as they are not voltage-dependent and are not blocked by 1,4-dihydropyridine calcium channel antagonists.

Thrombin-activated  $\text{Ca}^{2+}$  channels derived from platelet membranes were incorporated into artificial phospholipid bilayers and single channel currents recorded. Other types of ionic channels ( $\text{K}^+$ ,  $\text{Cl}^-$ ) were incorporated into planar lipid bilayers. However, neither in symmetrical (150 mM  $\text{BaCl}_2$  on *cis* side and *trans* side,  $n = 30$ ) nor asymmetrical (100 mM  $\text{BaCl}_2$  on *cis* side and 50 mM *trans* side,  $n = 5$ )  $\text{BaCl}_2$  solutions could single channel activity be detected in membranes from unstimulated platelets. These results suggest that the activity of the  $\text{Ca}^{2+}$  channels reported below is dependent on thrombin stimulation.

To identify calcium channels in platelet membranes, a high divalent cation concentration gradient, similar to that used for the study of voltage-dependent calcium channels, was employed. Barium was used as the charge carrier to facilitate the comparison of platelet calcium channels with single voltage-dependent calcium channels from other tissues. Figure 1a presents the current amplitude distribution in asymmetric solutions. A corresponding original record is shown in Fig. 1b. The holding potential was 0 mV and thus the positive current measured is carried by  $\text{Ba}^{2+}$ , indicating a selectivity for divalent cations. Figure 1a shows that the conductance of the open channel presented in Fig. 1b is composed of two main amplitudes with values very close to each other.

The experiments shown in Fig. 2 examine the effect of membrane potential on channel properties. Figure 2a and 2b show representative single channel recordings in symmetric and asymmetric solutions and Fig. 2c shows the single channel current-voltage relationship for the same channel under identical conditions. The single channel conductances in symmetric and asymmetric solutions were 10 pS and 8 pS, respectively.

The current-voltage relationship was linear in both sets of solutions, except that in the asymmetrical configuration the curve shape flattened at the negative holding potentials. For the purpose of calculating the selectivity of the channel for  $\text{Ba}^{2+}$  over  $\text{Na}^+$ , we took the membrane potential  $-70$  mV for zero current flow to be the reversal potential and calculated a relative selectivity for  $\text{Ba}^{2+}$  over  $\text{Na}^+$  of 35 according to a modified Goldman-Hodgkin-Katz equation<sup>4</sup>.

Open- and closed-time histograms made at zero holding potential could be fitted by two exponential curves, suggesting the existence of two open ( $[T_{o1}] = 0.5$  s and  $[T_{o2}] = 5.6$  s) and two closed ( $[T_{c1}] = 0.2$  s and  $[T_{c2}] = 2.0$  s) states for these divalent cation channels (data not shown). The probability of the channel being open at zero and  $+30$  mV holding potentials is 0.74 and 0.68, respectively, indicating a lack of voltage regulation of the channels' open state probabilities.

Divalent cations like  $\text{Cd}^{2+}$ ,  $\text{Co}^{2+}$  and  $\text{Ni}^{2+}$  are known to be effective  $\text{Ca}^{2+}$ -channel blockers. The effect of 20 mM  $\text{Ni}^{2+}$  on the channel activity in asymmetrical solutions at a holding potential of  $+50$  mV was investigated and the distribution of channel amplitudes in the closed and open state is shown in Fig. 3.  $\text{NiCl}_2$  diminished the mean unitary currents of the open state. A clear inhibition of the current could be demonstrated only at higher holding potentials ( $\geq +50$  mV). A reduction in the mean single channel current is consistent with the blocking events being too brief to resolve. An increased block at positive potentials is consistent with the blocking site being in the membrane field.

The 1,4-dihydropyridine calcium-channel blocker, nisoldipine (10  $\mu\text{M}$ ), had no effect on the mean unitary current ( $0.21 \pm 0.06$  pA and  $0.20 \pm 0.05$  pA in the absence and presence of nisoldipine, respectively, at a holding potential of  $+15$  mV) or the probability of the channel being open.

The effects of  $\text{Ni}^{2+}$  and nisoldipine on calcium channel activity are in agreement with earlier studies. Both  $\text{Cd}^{2+}$  and  $\text{Ni}^{2+}$  have been shown to have a potent inhibitory effect on thrombin-stimulated  $\text{Ca}^{2+}$  uptake and aggregation in mammalian platelets<sup>5-7</sup>, whereas organic calcium channel blockers, especially 1,4-dihydropyridine analogues, are ineffective<sup>7,8</sup>.

By incorporating elements from isolated human platelet mem-

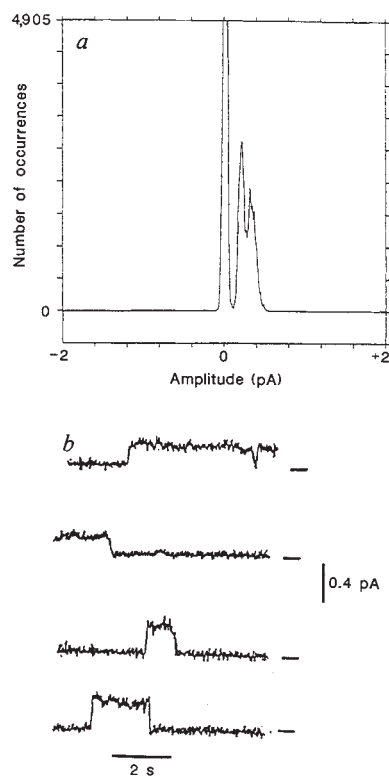
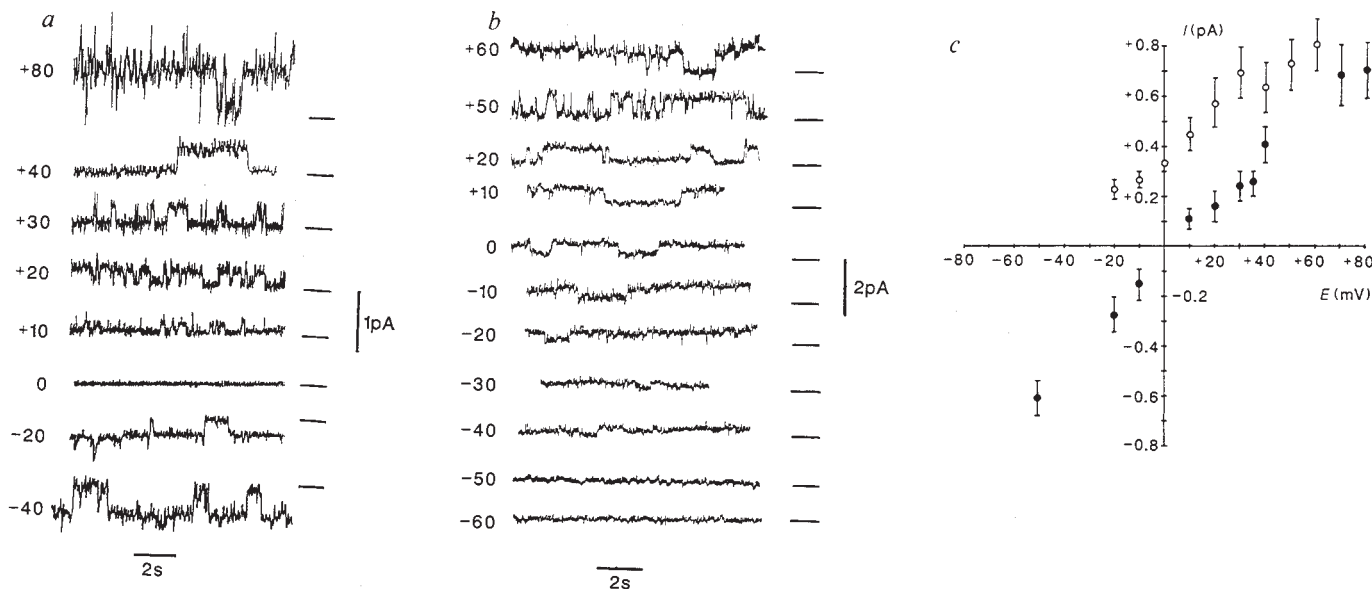


Fig. 1 a, Distribution of current amplitudes in the open and closed state of channels. The peak on the left represents the distribution of current when the channel was closed; the peak on the right represents the open state of the channel. The difference between the peaks represents the mean unitary current. The area under the peaks shows the proportional open/closed time. In this preparation the open state had two very close conductance levels (see text). The holding potential is 0 mV in asymmetrical (100 mM  $\text{BaCl}_2$  *cis* side and 50 mM  $\text{NaCl}$  *trans* side) solutions. b, Original recording of the single channel activity under the same conditions as above, filtered at 70 Hz (8-pole Bessel filter). The bars beside the recordings indicate the state when the channel is closed.

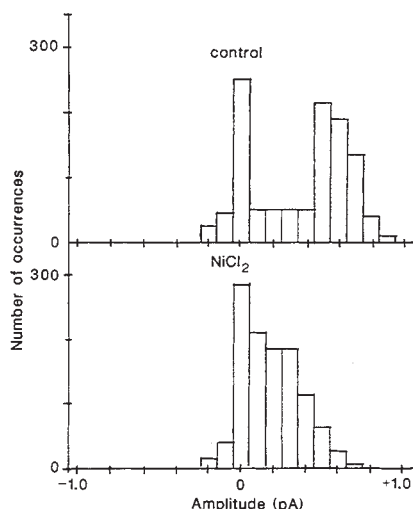
**Methods.** Platelets were isolated from human blood as described before<sup>17</sup>, and incubated in physiological salt solution for 1 h before adding thrombin ( $0.2 \text{ U ml}^{-1}$ ) for 2 min. After stimulation with thrombin, the platelets were frozen in liquid  $\text{N}_2$  and subsequently thawed on ice for the preparation of membranes at  $4^\circ\text{C}$ . The membranes were centrifuged at 40,000g for 10 min, resuspended with 20 mM HEPES and centrifuged at 40,000g for 10 min. The resuspension and centrifugation was repeated once more before resuspending the membranes in sucrose solution (0.4 M sucrose, 20 mM HEPES and 1% bovine serum albumin). Platelet membranes prepared in this way have been shown to have adenylate cyclase and calmodulin-dependent  $\text{Ca}^{2+}$ -ATPase activities<sup>18,19</sup>, indicating that they contain plasma membranes. The preparation was preserved at  $-75^\circ\text{C}$ . Planar lipid bilayers (30 mg  $\text{ml}^{-1}$  phosphatidylethanolamine and 25 mg  $\text{ml}^{-1}$  phosphatidylserine from bovine brain; Avanti Polar Lipids) were formed over a 100–250  $\mu\text{m}$  hole in a lexan partition; bilayer electronics were standard<sup>20</sup>. Platelet membranes were added to the *cis* side of the bilayer, the opposite *trans* side was grounded and a holding potential was applied to the *cis* side. The current direction was defined as positive when positive charge moved from the *cis* side to the *trans* side. Barium was used instead of  $\text{Ca}^{2+}$  as it has been shown to be more permeant through  $\text{Ca}^{2+}$  channels on other tissues<sup>20-22</sup>.

branes into planar phospholipid bilayers, we have obtained direct evidence for divalent cation selective channels in thrombin-stimulated platelets. This type of receptor-operated  $\text{Ca}^{2+}$  channel was originally proposed to explain agonist-induced dihydropyridine-insensitive  $\text{Ca}^{2+}$  entry in smooth muscle<sup>9,10</sup>. Benham and Tsien recorded single ATP-activated channel currents in smooth muscle which had 3:1 selectivity for  $\text{Ca}^{2+}$  over





**Fig. 2** *a, b*, Original recordings of single channel activity from thrombin-activated platelet membranes incorporated in a phospholipid bilayer separating symmetrical (*a*, 150 mM BaCl<sub>2</sub> on *cis* and *trans* side) and asymmetrical (*b*, 100 mM BaCl<sub>2</sub> on *cis* side and 50 mM NaCl on *trans* side) BaCl<sub>2</sub> solutions. All solutions used for single channel activity measurements contained in addition 10 mM HEPES at pH 7.0. Recordings are filtered at 70 Hz (8-pole Bessel filter). The holding potentials are indicated beside the recordings in mV. Bars beside the recordings indicate the state when the channel current is zero. *c*, Single channel current-voltage relations for Ba<sup>2+</sup> in symmetrical (●) and asymmetrical (○) solutions (the same constitutions as above). Each point represents the mean of all experiments with standard deviations. The slope calculated from data (in asymmetrical solution the data between 0 mV and +60 mV only was used) by least-squares fit gave single channel conductances with symmetrical and asymmetrical solutions  $10 \pm 3$  ( $n = 3$ ) and  $8 \pm 4$  ( $n = 5$ ), respectively.



**Fig. 3** Distribution of current amplitudes in the closed and open state of the channel in the absence and presence of 20 mM NiCl<sub>2</sub>. The histogram represents data from recordings made over a time period of 20 s at the holding potential of +50 mV, in asymmetrical solutions (100 mM BaCl<sub>2</sub> on *cis* side and 50 mM NaCl on *trans* side); each current amplitude was measured over 20 ms intervals. The mean opening single channel currents in the absence and presence of NiCl<sub>2</sub> (20 mM) were  $0.67 \pm 0.05$  pA ( $n = 100$ ) and  $0.43 \pm 0.05$  pA ( $n = 100$ ), respectively.

Na<sup>+</sup> (ref. 11). Receptor- and stretch-activated channels which promote Ca<sup>2+</sup> entry into endothelial cells also appear to be rather non-selective cation channels<sup>12</sup>. The present study thus reports on the first example of voltage-insensitive receptor-operated Ca<sup>2+</sup> channels.

The requirement of intact platelets for the activation of channels by thrombin implies that channel activation is more complex

than direct ligand-induced conformational change of a channel molecule, as proposed for the nicotinic acetylcholine receptor<sup>13</sup>. Coupling of the thrombin receptor to channel opening may instead be effected by the generation of second messengers, as suggested by the 200 ms delay between thrombin application and Ca<sup>2+</sup> entry in platelets<sup>14</sup>. Inositol-(1,3,4)triphosphate has been proposed for such intermediate messenger function in neutrophils<sup>15</sup> and inositol-(1,3,4,5)tetrakisphosphate in sea urchin eggs<sup>16</sup>.

Future use of the planar phospholipid bilayer technique should prove valuable in identifying the mechanism of receptor-operated calcium channel activation in platelet membranes.

This work has been supported by Fritz Thyssen Foundation, Swiss National Fund, NIH and American Heart Association. M.T.N. is an Established Investigator of the American Heart Association.

Received 16 May; accepted 14 June 1988.

1. Rink, T. J. & Hallam, T. J. *TIBS* **9**, 215-219 (1984).
2. Hallam, T. J. & Rink, T. J. *J. Physiol.* **368**, 131-146 (1985).
3. Doyle, V. M. & Ruegg, U. T. *Biochem. biophys. Res. Commun.* **127**, 161-167 (1985).
4. Fatt, P. & Ginsborg, B. L. *J. Physiol., Lond.* **142**, 516-543 (1958).
5. Hallam, T. J. & Rink, T. J. *FEBS Lett.* **186**, 175-179 (1985).
6. Jy, W. & Haynes, D. H. *Biochim. biophys. Acta* **929**, 88-102 (1987).
7. Blache, D., Ciavatti, M. & Ojeda, C. *J. Physiol.* **358**, 68P (1985).
8. Rink, T. J. *Agents and Action* **20**, (Suppl.), 147-169 (1986).
9. van Breemen, C., Aaronson, P. & Loutzenhiser, R. *Pharmac. Rev.* **30**, 167-208 (1979).
10. Bolton, T. B. *Physiol. Rev.* **59**, 607-718 (1979).
11. Benham, C. D. & Tsien, R. W. *Nature* **328**, 275-278 (1987).
12. Lansman, J. B., Hallam, T. J. & Rink, T. J. *Nature* **325**, 811-813 (1987).
13. Kistler, J., Stroud, R. M., Klymkowski, M. W., Lalancette, R. A. & Fairclough, R. H. *Biophys. J.* **37**, 371-383 (1982).
14. Sage, S. O. & Rink, T. J. *Biochem. biophys. Res. Commun.* **136**, 1124-1129 (1986).
15. Kuno, M. & Gardner, P. *Nature* **326**, 301-306 (1987).
16. Houslay, M. D. *TIBS* **12**, 133-134 (1987).
17. Resink, T. J. & Bühler, F. R. *J. Receptor Res.* **4**, 537-556 (1984).
18. Resink, T. J., Dimitrov, D., Stucki, S. & Bühler, F. R. *Biochem. biophys. Res. Commun.* **138**, 215-223 (1986).
19. Resink, T. J., Stucki, S., Grigorian, G. Y., Zschauer, A. & Bühler, F. R. *Eur. J. Biochem.* **154**, 451-456 (1986).
20. Nelson, M. T. *Ion Channel Reconstitution* (ed. Miller, C.) 507-523 (Plenum, New York).
21. Nelson, M. T. *J. gen. Physiol.* **87**, 201-222 (1986).
22. Nelson, M. T., French, R. J. & Krueger, B. K. *Nature* **308**, 77-80 (1984).

## Antigenic peptides recognized by T lymphocytes from AIDS viral envelope-immune humans

Jay A. Berzofsky\*, Armand Bensussan†, Kemp B. Cease\*††, Jean François Bourget‡, Remi Cheynier‡, Zirimwabagangabo Lurhuma§, Jean-Jacques Salaün||, Robert C. Gallo¶, Gene M. Shearer# & Daniel Zagury‡

\* Metabolism Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

† Hôpital St Louis, INSERM U93, 75475 Paris, France

‡ Physiologie Cellulaire, Université Pierre et Marie Curie, Paris, France

§ Cliniques Universitaires de Kinshasa, Kinshasa, Zaire

|| Institut National de Recherches Biomédicales, Kinshasa, Zaire

¶ Laboratory of Tumor Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892 USA

# Immunology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

T-lymphocyte immunity is likely to be an important component of the immune defence against the AIDS virus, because helper T cells are necessary for the antibody response as well as the cytotoxic response. We have previously predicted two antigenic sites of the viral envelope protein gp120 likely to be recognized by T lymphocytes, based on their ability to fold as amphipathic helices, and have demonstrated that these are recognized by T cells of mice immunized with gp120 (ref. 1). A peptide corresponding to one of these sites can also induce immunity in mice to the whole gp120 protein<sup>1</sup>. Because many clinically healthy seropositive blood donors have already lost their T-cell proliferative response to specific antigen<sup>2,3</sup>, we tested the response to these synthetic peptides of lymphocytes from 14 healthy human volunteers who had been immunized with a recombinant vaccinia virus containing the AIDS viral envelope gene and boosted with a recombinant fragment. Eight of the 14 responded to one peptide, and four to the other peptide, not included in the boost. These antigenic sites recognized by human T cells may be useful components of a vaccine against AIDS. We also found a correlation between boosting with antigen-antibody complexes (compared to free antigen) and higher stimulation indices, suggesting a more effective method of immunization.

Because T lymphocyte immunity is likely to be important in preventing or controlling an infection with the human immunodeficiency virus (HIV), and because immunization with killed preparations of whole HIV or subfractions thereof may not be safe or practical, it is important to identify those regions of the HIV proteins that can induce T-cell immunity so that these can be incorporated into vaccine preparations. Our approach to this problem has been to predict antigenic sites on the basis of their ability to fold as amphipathic helices<sup>4-7</sup>, synthesize the corresponding peptides, and assess their function in mice and subsequently in primates. The first two sites we identified in the envelope of the HTLV-III<sub>B</sub> isolate of HIV, called envT1 and envT2, corresponded to amino acid residues 428-443 (Lys-Gln-Ile-Ile-Asn-Met-Trp-Gln-Glu-Val-Gly-Lys-Ala-Met-Tyr-Ala) and 112-114 (His-Glu-Asp-Ile-Ile-Ser-Leu-Trp-Asp-Gln-Ser-Leu-Lys) of the gp120 protein, respectively. These were shown to elicit class II-MHC molecule restricted T-cell immunity in mice<sup>1</sup>. To assess the ability of these peptides to be recognized by HIV-immune T cells in humans, we first examined peripheral blood lymphocytes from naturally infected, seropositive blood donors who were still clinically healthy. However, because a significant fraction of these had already lost their *in*

*vitro* response to control antigens such as influenza or tetanus toxoid (see also refs 2, 3) we could not adequately assess their response. Therefore, we took advantage of the opportunity to perform a retrospective study of a group of healthy, initially seronegative volunteers who had been immunized, with the approval of the French and Zairian regulatory authorities, with a recombinant vaccinia virus<sup>8</sup> expressing the gene for the HIV envelope protein gp160 of the HTLVIII<sub>B</sub>, BH10, isolate<sup>9,10</sup>. All the volunteers were immune as judged by antibodies to the envelope protein. Some had also been recently boosted with a recombinant fragment of the envelope representing approximately the carboxy-terminal 40% of the gp120 portion of the molecule, and containing the putative epitope 428-443 (envT1) but not 112-124 (envT2). Peripheral blood mononuclear cells from these donors were tested *in vitro* for their ability to proliferate in response to various control antigens and the synthetic peptides envT1 and envT2 corresponding to these epitopes.

The donor most thoroughly studied, D.Z., had been immunized six months previously with the recombinant vaccinia virus and boosted three months later with a slow-drip infusion of

**Table 1** T-cell proliferation to peptide envT1 in peripheral blood of volunteers immunized with gp160-vaccinia recombinant: effect of boosting protocol

| Expt | Boost protocol     | Donor | Stimulation dose <i>in vitro</i> | Stimulation index* | P value† |
|------|--------------------|-------|----------------------------------|--------------------|----------|
| 1    | protein + antibody | 1     | 4 µM                             | 4.3                | <0.05    |
|      |                    | 2     | 4 µM                             | 25.8               | 0.005    |
|      |                    | 3     | 4 µM                             | 41.7               | <0.025   |
|      | protein only       | 4     | 4 µM                             | 1.8                | ind‡     |
|      |                    | 5     | 4 µM                             | 1.1                | NS       |
|      |                    | 6     | 4 µM                             | 9.1                | <0.025   |
|      |                    | 7     | 4 µM                             | 2.4                | ind‡     |
|      | none               | 8     | 4 µM                             | 0.5                | NS       |
|      |                    | 9     | 4 µM                             | 1.2                | NS       |
| 2    | protein + antibody | 10    | 1 µM                             | 2.8                | <0.1     |
|      |                    | 11    | 1 µM                             | 5.1                | ind.‡    |
|      |                    | 12    | 1 µM                             | 15.7               | <0.025   |
|      | protein only       | 13    | 1 µM                             | 1.7                | <0.05    |
|      | none               | 14    | 1 µM                             | 0.5                | NS       |
| 3    | none§ (controls)   | G201  | 4 µM                             | 0.33               | NS       |
|      |                    | G202  | 4 µM                             | 0.59               | NS       |
|      |                    | G204  | 4 µM                             | 0.67               | NS       |

All donors except the controls in expt 3 had been immunized ~six months previously with envelope-recombinant vaccinia virus. Where indicated, some had also been boosted 44 days previously with soluble antigen or antigen-antibody complexes as shown (see text). Donor 9, listed as not boosted at this time, had been boosted 116 days previously. As expected from the time course in Fig. 1, there was no response after this long interval after the boost. Because of limited cell numbers, and limited typing sera for African populations, partial HLA class II typing data could be obtained on only six of the donors, as follows: donor 2 (DR 2/4, DQ untestable); donor 6 (DR 5/6, DQ 1/3); donor 8 (DR 3/7, DQ untestable); donor 10 (DR 2/7; DQ 1/-); donor 12 (DR 5/-, DQ untestable); donor 13 (DR2/-, DQ untestable).

\* Ratio of experimental mean c.p.m. <sup>3</sup>H-thymidine incorporation with peptide over control mean c.p.m. without antigen. Background proliferation without antigen ranged from 558 to 7,803 c.p.m., with most donors in the range of 1,000-5,000 c.p.m. All donors showed strong proliferative responses to positive control antigens PPD and/or streptococcal enzyme varidase, with stimulation indices >10 for one or both of these.

† Probability based on a one-tailed Student's *t* test for replicates in the presence of peptide compared to control replicates without antigen. Values of <0.05 were taken as significant, but values just barely >0.05 are listed as <0.1 to indicate borderline significance. NS, not significant.

‡ Indeterminate due to shortage of cells requiring single value of either experimental or control well.

§ Healthy seronegative donors who were not immunized. All had proliferative responses to influenza with stimulation indices of 13.6, 26.8 and 14.8, respectively.

†† Present address: Department of Internal Medicine, University of Michigan School of Medicine, Ann Arbor, Michigan 48109, USA.

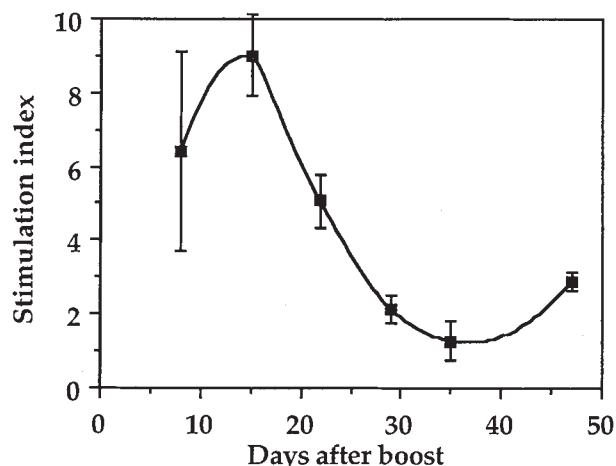


$3.5 \times 10^7$  paraformaldehyde-fixed autologous cells that had been infected with the recombinant vaccinia virus *in vitro*<sup>10</sup>. This donor responded initially to peptide envT2 (112-124) (22,561 c.p.m. <sup>3</sup>H-thymidine incorporation compared to a control with medium alone of 1,337 c.p.m.) and to gp120 (29,908 c.p.m.), but only marginally to envT1 (428-443) (2,747 c.p.m.). After a boost<sup>10</sup> with gp160, however, his peripheral blood lymphocytes proliferated strongly to the envT1 peptide (428-443) with the time course shown in Fig. 1. The response peaked at two weeks after the boost and gradually fell to a lower level of two- to threefold over background, still detectable 4-7 weeks after the boost. Thus, at least this one human immunized with the envelope protein in the form of a vaccinia virus recombinant developed significant T-cell immunity specific for both epitopes, and the immunity to the envT1 site (428-443) could be boosted specifically by immunization with a large recombinant protein containing that site.

To determine whether responses to these sites were frequent in humans immunized with the HIV envelope-expressing recombinant vaccinia virus, we studied a larger panel of 14 healthy human volunteers that had been similarly immunized six months previously. Blood from these donors was available for study 44 days after some of them had been boosted with the recombinant fragment of gp120 containing the envT1 site 428-443. Two different boosting protocols had been employed, giving us the opportunity to compare these protocols for boosting T-cell immunity. Some had been re-immunized with the recombinant carboxyl-terminal 40% of gp120 alone<sup>10</sup>, whereas others received a mixture of this fragment and anti-envelope antibodies (1 mg antigen with 50  $\mu$ l purified monoclonal antibody 9284, developed by Biotech Research Laboratories and marketed by Dupont, specific for gp120, injected intramuscularly without adjuvant). The results (Table 1) indicate that 8 of the 14 gave proliferative responses more than twice background. Of these, five were significantly higher than the medium controls and one was close to the level of significance. Two gave inadequate cell numbers to perform replicate cultures and thus significance could not be determined. The two best responders, donors 2 and 3, gave proliferative responses of 119,526 and 58,463 c.p.m. above background, respectively. Altogether, 8 of the 11 individuals boosted responded, despite the long interval after the booster immunization. None of three normal unimmunized seronegative volunteer blood donors responded to either peptide (Table 1). Thus, T cells from a large fraction of immunized humans recognize the peptide 428-443 (envT1) as an antigenic site.

Even more striking is the apparent correlation between the type of booster immunization and the magnitude of response. Of the group tested at 4  $\mu$ M peptide concentration *in vitro* (expt 1), all three boosted with the mixture of protein plus antibody responded significantly, whereas only two of the four boosted with protein alone responded and neither of the two unboosted individuals responded. Moreover, the geometric mean stimulation index of the three boosted with the mixture (16.7) was significantly higher than that of the group boosted with protein alone (mean 2.56,  $P < 0.05$ ), and that of the unboosted group (mean 0.77,  $P < 0.025$ ). Because the groups boosted with protein only or with nothing in expt 2 contained only one member each, the comparisons could not be made statistically, but the trend is clearly the same (Table 1). All three boosted with the mixture responded with indices  $> 2$ , but neither of the others gave responses  $> 2$ , although the one given protein only was statistically higher than background. We conclude that boosting with antigen-antibody complexes of envelope protein and anti-envelope antibody improves T-cell proliferative responses in humans.

Perhaps because none of the donors had been boosted recently with material containing the envT2 site (112-124), only 4 of the 14 (donors 1, 4, 10 and 11) gave proliferative responses to the envT2 peptide greater than twofold over background. Neverthe-



**Fig. 1** Kinetics of peripheral blood T-cell proliferative response to peptide envT1 (gp160 residues 428-443) as a function of time after booster immunization. Fresh heparinized blood was drawn from donor D.Z. at the times shown after he had been boosted<sup>10</sup> with 100  $\mu$ g of gp160 purified from cells infected with the recombinant vaccinia virus carrying the HIV *env* gene, prepared as described by Chakrabarti *et al.*<sup>5</sup>. The peripheral blood mononuclear cells were separated on a layer of Ficoll-Hypaque and resuspended in RPMI 1640 medium with 10% normal human serum (decomplemented), 2 mM L-glutamine, 100 U ml<sup>-1</sup> penicillin, and 100  $\mu$ g ml<sup>-1</sup> streptomycin, and cultured at  $10^5$  cells in 0.2 ml per well in 96-well flat-bottom microtitre plates (Falcon 3072) with 3  $\mu$ M peptide envT1 for six days at 37 °C, 5% CO<sub>2</sub>. During the last 8 h, 1  $\mu$ Ci of [<sup>3</sup>H]-thymidine was added per well to measure the incorporation into DNA as a measure of cell proliferation. The ordinate shows the stimulation index, the ratio of geometric experimental mean c.p.m. to the geometric mean control c.p.m. (without antigen) of triplicate or duplicate cultures. Control cultures ranged from 300 c.p.m. to 1,434 c.p.m., except for the last time point at which the control mean was 6,235 c.p.m. The error bars represent standard error of the mean.

less, these results, taken together with the results on donor D.Z. obtained earlier, indicate that at least a sizeable minority of humans also respond to the envT2 site. The fraction positive might have been larger if we had been able to test their responses to peptide closer to the time of immunization or boosting with material containing the envT2 site. However, constraints on work with human subjects prevented additional immunizations for these experiments.

We would expect that the proliferative responses to these peptides should be restricted to class II MHC antigens, and that the response to individual epitopes, as represented by these peptides, might be under HLA-linked *Ir* gene control. Thus, it is possible that some of those who failed to respond did so because they had a low responder class II HLA type for that epitope, although as responsiveness is usually dominant, unresponsiveness requires two alleles of the low responder type. Because of the limited blood specimens available from Zaire, although we were able to obtain HLA class I (A, B, C) typing data on 12 of the donors, we only had enough cells available for partial class II (DR/DQ/DP) typing on six of the donors (see legend to Table 1), and we could not obtain Epstein-Barr virus-transformed B-cell lines from frozen cells to do further typing. The donors, however, are from an outbred population and nearly all unrelated; as expected, the HLA class I typing data obtained indicated a heterogeneous group, with no apparent correlation between HLA type and response or nonresponse. Some of the same HLA-A, B and C alleles were seen in individuals who responded and ones who did not. Because of linkage disequilibrium, one would expect the class II types to be equally heterogeneous, and this is borne out by the DR typing on six of the individuals (Table 1, legend). Thus, even if class II typing on all of the blood donors were available, this

study would be very unlikely to shed light on the question of HLA-linked genetic control of the response to these peptides in humans. The probability of being able to detect a statistically significant correlation between responsiveness and HLA-DR, DQ or DP type with this size sample is extremely small. Thus, at this point we cannot draw any conclusions about HLA linkage, and the results of additional HLA typing on this small donor population would not change any of our other conclusions. In this type of study of a small group of humans, one must consider HLA type as one of the many variables that differ among outbred individuals that cannot be controlled. The fact that the differences between groups were statistically significant despite these variables and despite the small sample size, makes the results that much more striking. If the differences between groups had been subtle, we would not have detected them with this size of sample.

These results demonstrate that a sizeable fraction of humans immunized with the HIV envelope protein (in the form of a recombinant vaccinia virus expressing the gp160 gene) recognize the T-cell antigenic sites envT1 (428–443) or envT2 (112–124) or both. Thus, these sites, first described in murine T-cell responses to the HIV envelope<sup>1</sup>, are also immunogenic in humans. Therefore, these sites may be important to include in a vaccine intended to develop helper T-cell immunity to HIV, which we expect to be necessary to obtain either an anamnestic antibody response or a cytotoxic T-cell response. The proliferating T cells may also include those that directly mediate delayed hypersensitivity, although the role of this type of immune response in HIV infection is unknown. Knowledge of the location of these sites should be useful in designing a vaccine. Moreover, as these two T-cell antigenic sites are in relatively conserved regions of the envelope sequence, and are active in the native envelope protein, they should be effective in eliciting immunity crossreactive with a large number of viral isolates (group-specific rather than type-specific immunity). Group-specific T-cell immunity to gp120 has been reported in goats<sup>11</sup>.

It is of interest that the site 428–443 closely overlaps a portion of the envelope sequence suggested by Lasky *et al.*<sup>12</sup> to be part of the site for binding to the CD4 cell surface receptor. As this function probably requires more than just this short segment of the protein, and may depend on the three-dimensional conformation (tertiary structure) of the molecule, it is unlikely that the envT1 peptide (428–443) itself will bind to CD4. Except for the fact that this peptide comes from a relatively conserved, nonglycosylated portion of gp120, as one would also expect for the CD4 binding site, we cannot at this time account for this coincidence.

The use of antigen-antibody complexes to enhance the immunogenicity of protein antigens has been studied in experimental animals intermittently for two decades<sup>13–17</sup>. Doses of antigen too low to be immunogenic alone, in adjuvant, were found to be immunogenic when complexed with specific antibody. Immune complexes are probably more efficiently taken up by macrophages and other antigen-presenting cells via their Fc receptors, after which the complexes are processed into fragments and such fragments associate with class II major histocompatibility molecules on the surface of these antigen-presenting cells where they can stimulate antigen-specific T cells. We are not aware that this technique had been used before in immunizing humans for T-cell immunity. The current results strongly suggest that this technique may be especially effective in inducing helper T-cell immunity to specific T-cell antigenic sites in humans as well.

7. Berzofsky, J. A. *et al. Immun. Rev.* **98**, 9–52 (1987).
8. Chakrabarti, S., Robert-Guroff, M., Wong-Staal, F., Gallo, R. C. & Moss, B. *Nature* **320**, 535–537 (1986).
9. Zagury, D. *et al. Nature* **326**, 249–250 (1987).
10. Zagury, D. *et al. Nature* **332**, 728–731 (1988).
11. Krohn, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 4994–4998 (1987).
12. Lasky, L. A. *et al. Cell* **50**, 975–985 (1987).
13. Terres, G. & Wolins, W. *J. Immun.* **86**, 361 (1961).
14. Möller, G. & Wigzell, H. *J. exp. Med.* **121**, 969–989 (1965).
15. Uhr, J. W. & Möller, G. *Adv. Immun.* **8**, 81–127 (1968).
16. Celis, E. & Chang, T. W. *Science* **224**, 297–299 (1984).
17. Chang, T. W. *Immun. Today* **6**, 245–249 (1985).

## Membrane guanylate cyclase is a cell-surface receptor with homology to protein kinases

Sujay Singh\*†, David G. Lowe§, David S. Thorpe‡, Henry Rodriguez¶, Wun-Jing Kuang§, Lawrence J. Dangott\*‡, Michael Chinkers\*‡, David V. Goeddel§ & David L. Garbers\*‡‡

\* The Howard Hughes Medical Institute, Departments of † Pharmacology and ‡ Molecular Physiology and Biophysics, Vanderbilt Medical Center, Nashville, Tennessee 37232, USA § Department of Molecular Biology and ¶ Department of Developmental Biology Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

Guanylate cyclase has been strongly implicated as a cell-surface receptor on spermatozoa for a chemotactic peptide<sup>1</sup>, and on various other cells as a receptor for atrial natriuretic peptides<sup>2–4</sup>. Resact (Cys-Val-Thr-Gly-Ala-Pro-Gly-Cys-Val-Gly-Gly-Arg-Leu-NH<sub>2</sub>), the chemotactic peptide released by sea urchin *Arbacia punctulata* eggs, is specifically crosslinked to *A. punctulata* spermatozoan guanylate cyclase<sup>1</sup>. After the binding of the peptide the state of guanylate cyclase phosphorylation modulates enzyme activity<sup>1,5,6</sup>. We report here that the deduced amino-acid sequence of the spermatozoan membrane form of guanylate cyclase predicts an intrinsic membrane protein of 986 amino acids with an amino-

**Table 1** Immunoprecipitation of guanylate cyclase activity with antisera to synthetic peptide and purified 160K protein

| Rabbit serum            | pmol cGMP‡ |
|-------------------------|------------|
| Peptide antiserum*      | 1,000      |
| Preimmune               | 230        |
| 160K protein antiserum† | 7,050      |
| Preimmune               | 110        |

\* For immunoprecipitation of guanylate cyclase with pre-immune or synthetic peptide antiserum (see text) 100 µl of an approximate 25% (w/v) suspension of intact spermatozoa were incubated with 20 µl of either preimmune or immune serum for 30 min at 2 °C. Nine volumes of homogenization buffer (Fig. 1 legend) were added, cells were disrupted in a dounce homogenizer, and the homogenate was clarified by centrifugation at 13,000g for 10 min. An 0.5-ml aliquot was incubated with formalin fixed *S. aureus* cells (BRL) for 30 min at 2 °C. Immune complexes were recovered by centrifugation and washed twice in 25 mM Tris pH 7.5, 150 mM NaCl, 0.1% Nonidet P-40 and three times in 25 mM Tris pH 7.5, 150 mM NaCl.

† Protein A affinity purified IgG from preimmune serum or antiserum to purified 160K protein were incubated with clarified sperm homogenate, prepared as described above, for 1.5 h at 2 °C. Immune complexes were recovered with formalin fixed *S. aureus* cells and washed as described above.

‡ To determine guanylate cyclase activity, *S. aureus* pellets were resuspended in homogenization buffer (Fig. 1 legend) with 0.2 mM [ $\alpha$ -<sup>32</sup>P]GTP (for peptide antiserum) or 0.1 mM [ $\alpha$ -<sup>32</sup>P]GTP (protein antiserum), and 5 mM MnCl<sub>2</sub>, incubated for 10 min at 30 °C, and formed [<sup>32</sup>P]cGMP was measured as previously described<sup>25</sup>. Results are expressed as the total amount of cGMP formed in each reaction.

Received 18 March; accepted 23 June 1988.

1. Cease, K. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 4249–4253 (1987).
2. Shearer, G. M. *et al. J. Immun.* **137**, 2514–2521 (1986).
3. Wahren, B. *et al. J. Virol.* **61**, 2017–2023 (1987).
4. DeLisi, C. & Berzofsky, J. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7048–7052 (1985).
5. Spouge, J. L. *et al. J. Immun.* **138**, 204–212 (1987).
6. Margalit, H. *et al. J. Immun.* **138**, 2213–2229 (1987).



terminal signal sequence. A single transmembrane domain separates the protein into putative extracellular and cytoplasmic-catalytic domains. The cytoplasmic carboxyl-terminal 95 amino acids contain 20% serine, the likely regulatory sites for phosphorylation. Unexpectedly, the enzyme is homologous to the protein kinase family.

Guanylate cyclase appears to exist entirely as a membrane-bound glycoprotein in sea urchin spermatozoa<sup>7</sup>. The enzyme was purified to apparent homogeneity from the spermatozoa of the sea urchin *A. punctulata* by preparative electrophoresis (see Fig. 1 legend). Analysis of the purified protein by SDS-polyacrylamide gel electrophoresis (PAGE) and silver-staining demonstrated a single, major protein with a relative molecular mass of 160,000 (Fig. 1a). Antibody produced against a synthetic peptide, YNPTVINLDRGR (single-letter code<sup>8</sup>), derived from the amino-terminal peptide 1 sequence of the 160K protein (Fig. 1 legend), reacted with the purified protein on immunoblots (data not shown), and was used to probe blots of untreated or resact-treated sea-urchin sperm membrane extracts (Fig. 1b). Before the addition of resact, approximately 60–70% of the guanylate cyclase of the sperm membranes was in the phosphorylated, high molecular mass form<sup>5,9,10</sup>, and antibody bound to both the phosphorylated and dephosphorylated forms of the enzyme (Fig. 1b, lane 1). Resact is known to cause a change in the apparent molecular mass of guanylate cyclase through dephosphorylation of the enzyme<sup>9–11</sup>, and it also caused a complete shift in the mobility of the immunoreactive protein to the lower molecular mass form (Fig. 1b, lane 2). In addition, anti-

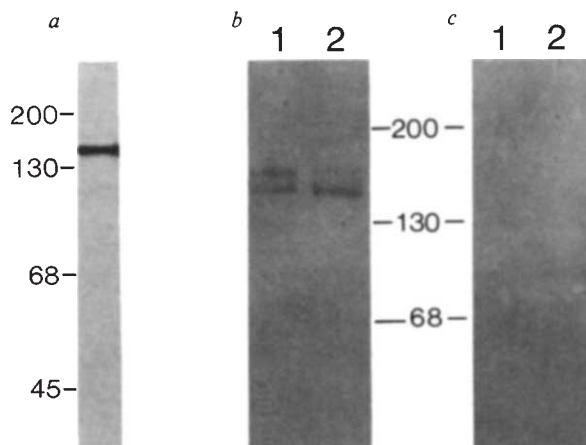
body to either the amino-terminal synthetic peptide, or purified 160K protein (Fig. 1 legend), immunoprecipitated guanylate cyclase activity from *A. punctulata* spermatozoa (Table 1). These data demonstrate that the 160K protein we have purified and for which we have obtained partial amino-acid sequence is the membrane-associated guanylate cyclase of sea-urchin spermatozoa.

A random-primed sea-urchin testis cDNA library in phage  $\lambda$ -ZAP was screened with two degenerate oligonucleotide probes based on the amino acid sequences of peptide 1 (probe A), and peptide 2 (probe B) (see Fig. 1 and Fig. 3 legends). Nucleotide sequence analysis (see below) of three clones, 31.1, 37.1 and 32.2 covered 3,444 base pairs (bp) and included the entire guanylate cyclase coding region of 2,958 bp and 5'- and 3'-untranslated sequences (Fig. 2a). Northern blot analysis of 5  $\mu$ g of poly(A)<sup>+</sup> RNA from *A. punctulata* testes was performed at high stringency using cDNA clone 31.1 as a probe (Fig. 2b). A single messenger RNA species of approximately 4.4 kilobases (kb) appeared to be relatively abundant. The entire coding sequence comprises 2,958 nucleotides and therefore approximately 1.5 kb of 5'- and 3'-flanking sequences are expected in the mRNA.

The 3,444-nucleotide sequence of cloned guanylate cyclase cDNAs (Fig. 3) contains an open reading frame of 986 codons. The assignment of the initiation methionine codon is based upon (1) two upstream in-frame stop codons, (2) the ATG at nucleotide position 141 is flanked by sequences that are conserved at translation initiation sites, an A in position -3 and a G in

**Fig. 1** Purification of the membrane form of guanylate cyclase and immunoblot analysis. *a*, Silver-stained gel of purified guanylate cyclase used for protein sequencing. Approximately 160 ng of purified protein were reduced, electrophoresed in a 12.5% polyacrylamide gel and silver-stained. *b*, The reaction of antibody to amino-terminal synthetic peptide with guanylate cyclase on immunoblots. Lane 1, plasma membrane preparation. Lane 2, plasma membrane preparation after treatment with resact. *c*, Same as *b* except the blot was probed with pre-immune rabbit serum. Molecular mass standards were myosin (200K),  $\beta$ -galactosidase (130K), bovine serum albumin (68K) and ovalbumin (45K).

**Methods.** Guanylate cyclase from sea-urchin spermatozoa was purified to homogeneity by preparative SDS-PAGE. In previous experiments we had established that the 160K protein was guanylate cyclase<sup>15</sup>. Twenty grams (wet weight) of intact spermatozoa were homogenized by sonication in 9 volumes (200 ml final volume) of a buffer containing 0.5% Lubrol PX, 35% glycerol, 25 mM 2[*N*-morpholino]ethanesulphonic acid (MES) (pH 6.2), 10 mM benzamidine hydrochloride, 10 mM NaF, and 5 mM dithiothreitol. The homogenate was centrifuged at 15,000 r.p.m. in a JA-20 rotor (Beckman) for 10 min at 2 °C. Proteins present in the supernatant fluid were denatured in a solution containing 100 mM Tris (pH 6.8), 1% SDS, and 40 mg ml<sup>-1</sup> bromophenol blue by heating at 57 °C for 10 min, and were electrophoresed through an SDS-polyacrylamide gel continuously eluted by gel buffer. Fractions containing the 160K protein (as determined by gel analysis) were pooled, concentrated using an Amicon Diaflo stir cell with a YM-30 membrane, and further purified by a passage through a Sephadex G-50 column equilibrated with 0.01% SDS and 0.1 M ammonium bicarbonate. Approximately 2 nmol of protein were obtained from an equivalent of 5 g of sperm. One nmol of the protein was digested with about 200 pmol of *S. aureus* V8 protease, and the digestion products were resolved by reverse phase HPLC (Vydac C18 column) pre-equilibrated with 5 mM ammonium formate (pH 6.5). The column was developed with a linear gradient (1% per min) of MeCN in 5 mM ammonium formate (pH 6.5) at a flow-rate of 1 ml min<sup>-1</sup>. Isolated peptides, or approximately 75 pmol of the intact purified protein, were subjected to automated Edman degradation on an Applied Biosystems 470A gas phase sequencer with an on-line model 120A PTH analyser. In other experiments, purified 160K protein was excised from SDS polyacrylamide gels and digested with *S. aureus* V8 protease. Peptides were subsequently eluted from gel slices and prepared for automated microsequencing. Amino-acid sequences of two *S. aureus* V8 protease-generated peptides, ATLXYNPTVINLDRGR (peptide 1) and LKTSINMAXVYIF (peptide 2), were used to design synthetic oligonucleotide probes for the isolation of DNA clones (see Fig. 3 legend). Seven additional amino-acid sequences were obtained from V8 protease-generated peptides: WEILHE (peptide 3), WEILHEE (peptide 4), IRDSQDYHSLE (peptide 5), YTLLE (peptide 6), ASDYLEQINQAYE (peptide 7), VQTR (peptide 8), and LQHE (peptide 9). Subsequent amino-terminal sequencing of the purified protein yielded the sequence ATLHYNPTVINLDRGR, which corresponded to peptide 1 listed above. For the preparation of antibodies a peptide (YNPTVINLDRGR) corresponding to the amino-terminal region of guanylate cyclase was synthesized by Peninsula Laboratories, and conjugated to bovine serum albumin with the bifunctional cross-linking agent disuccinimidyl suberate. The initial immunization of rabbits with the conjugated protein was performed with Freund's complete adjuvant. All subsequent booster injections were prepared with incomplete adjuvant. Antibodies against the purified guanylate cyclase were obtained by immunization of rabbits with protein electroeluted from polyacrylamide gel slices. Control or resact-treated plasma membrane fractions were prepared from *A. punctulata* spermatozoa according to previous methods<sup>5</sup>. Membrane proteins were solubilized and denatured in SDS sample buffer in the presence of reducing agent, electrophoresed through 7.5 to 15% polyacrylamide gradient slab gels, electroblotted overnight at 4 °C on to nitrocellulose membranes and probed with either the site-directed rabbit antiserum or pre-immune serum diluted 1:200 in phosphate buffered saline containing 5% non-fat dried milk as a blocking agent. The membranes were washed with buffered saline containing 0.1% sodium deoxycholate and 1.0% Tween-20. The immunoreactive bands were visualized with a Vectastain kit using horseradish peroxidase (Vector Labs) according to the manufacturer's protocol.



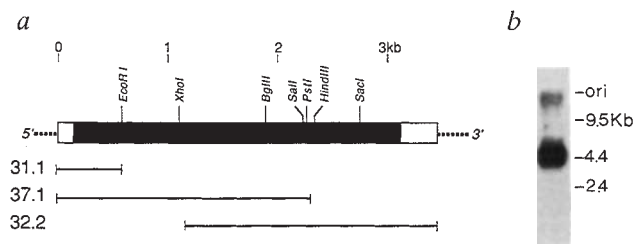
position +4 (ref. 12), and (3) the downstream 21-amino-acid sequence possesses all the features characteristic of a signal sequence<sup>13</sup>. Based on the amino-terminal protein sequence (Fig. 1 legend), we predict the signal peptide cleavage site to be after Ser-1, with the mature protein deduced to be composed of 965 amino acids with a calculated molecular mass of 106,150. The size differences between the purified enzymes (150,000–160,000)

**Methods.** A random-primed cDNA library was constructed in phage  $\lambda$ -ZAP (Stratagene) using poly(A)<sup>+</sup>RNA from *A. punctulata* testis and an Amersham cDNA synthesis kit. Approximately 200,000 independent clones were amplified in bacterial strain BB-4. Two 144-fold degenerate anti-sense oligonucleotide probes, probe A (5'-A/G/CGGCC(A/G)CG(A/G)TC-(C/G)AGGTTGATC(T/G/A)-GT(A/G/T)GGGTGTA-3'), and probe B (5'-GAAGATGTA(C/G)-ACTT(A/G)GCCAT(A/G)TTGAT(A/G)GA(A/G/T)GTCT-3'), were used to screen four hundred thousand clones from the amplified library<sup>26</sup>. Cloned cDNAs were rescued as Bluescript plasmids, utilizing the automatic excision process of  $\lambda$ -ZAP. To identify regions in the cloned cDNAs responsible for hybridization to probes A and B, *Sau3A*I fragments were shotgun cloned into M13mp18, screen by filter hybridization, and sequenced<sup>27</sup> to determine regions of identity with either probe.

and the estimated size from the deduced amino-acid sequence is possibly due to N-linked and/or O-linked glycosylation. There are four potential N-linked glycosylation sites at residues 6–8 (Asn-Pro-Thr), 164–166 (Asn-Ile-Thr), 340–342 (Asn-Gly-Thr), and 389–391 (Asn-Thr-Ser) (Fig. 3). The unambiguous determination of Asn in peptide 1 (Fig. 1 legend), however, suggests that it is not glycosylated. The phosphorylation state of the



**Fig. 2** Guanylate cyclase cDNA clones and Northern blot analysis of sea urchin testis mRNA. *a*, The bar at the top of the figure represents part of the guanylate cyclase mRNA, with the open and solid regions representing the untranslated and coding sequences, respectively. Untranslated flanking regions not cloned are represented by broken lines. Overlapping clones used in sequence determination are shown below the mRNA structure. *b*, Northern blot analysis of *A. punctulata* testis poly(A)+RNA (5 µg mRNA). Numbers on the right indicate the size of RNA markers (BRL, Inc.).



**Fig. 4** Guanylate cyclase amino-acid sequence similarity with protein kinases. Alignment of the sea-urchin guanylate cyclase predicted amino-acid sequence (GC) with the catalytic domains of the murine PDGF receptor<sup>21</sup> human c-fes<sup>22</sup>, and human c-mos<sup>23</sup>. 33 residues highly conserved in the protein kinases<sup>24</sup> are indicated by shading. Residues identical between GC and one or more of these three protein kinases are boxed. Dashes indicate gaps for optimal alignment, numbers in parentheses denote the number of amino acids in a variable region, and numbers at the end of GC and PDGF receptor sequences indicate the number of amino acids to the C-terminus.

|       |     |                                                                   |
|-------|-----|-------------------------------------------------------------------|
| GC    | 530 | T R E S E T N S Q F S M K S M V L S A I S V - - - - - I           |
| PDGFR | 566 | D Q L V I G R T L G S G A F G Q V V E A T A H G L S H S Q A T M   |
| c-fes | 562 | E D L V L G E Q I G R C N F G E V F S G R L R A - - - - - D N T   |
| c-mos | 58  | E Q V C L L Q R L G A G G F S S Y K A T Y R G - - - - - V         |
| GC    | 554 | S N E Q Q I F A T I G T Y R G T I C A I H A V H K N H I D L       |
| PDGFR | 598 | K V V M K S T A R S S - - - - - - - - - - - E                     |
| c-fes | 589 | L V V S C R E T L P P D - - - - - - - - - - - I                   |
| c-mos | 83  | P V A I K Q N K C T K N R L A - - - - - - - - - - S               |
| GC    | 586 | T R A V R T L K L M R D M - R H D N C P F I G A C I D - - -       |
| PDGFR | 613 | K Q A L M S E L K I M S H L G P H L N V N L L G A C T K G - -     |
| c-fes | 604 | K A K F L Q E A K I L K Q Y - S H P N V R L I G V C T Q K - -     |
| c-mos | 100 | R R S F W A E L N V A R L - - R H D N V R V V A A S T R T P A     |
| GC    | 614 | - - - R P H T I C I L M H Y C A K G S L Q D I M E N D D I K - - - |
| PDGFR | 643 | - - - G P - I Y I I T E Y C A K G S L Q D I M E N D D I K - - -   |
| c-fes | 633 | - - - Q P - I Y I I V M E L V Q G G D F L T F L R T E G A R - - - |
| c-mos | 130 | G S N S L - G T I I M E F G G N V T L H Q V I Y G A A G H (-15-)  |
| GC    | 640 | L D S M F L A S L I A D L V K G L V L H S S E I K S H G H L K     |
| PDGFR | 766 | L S Y T D L V G F S Y Q V A N G M D E L A S K - N C V H R D L A   |
| c-fes | 658 | L R V K T L L Q M V G D A A A G M E V I E S K - C C I H R D L A   |
| c-mos | 173 | L S L G K C L K Y S L D V V N G L L F L H S Q - S I V H L D L K   |
| GC    | 672 | S S M E V V D N R W V L Q T T D Y F G L H E F R K G Q K E D V D L |
| PDGFR | 797 | A R M L I C E G K L V K I C D F G L A R D I M R D S N Y I S K     |
| c-fes | 689 | A R M L V T E K N V L K I S D F G M S R E E A D G V Y A A S G     |
| c-mos | 204 | P A M L I S E Q D V C K I S D F G C S E K L E D L C F Q T P       |
| GC    | 704 | G E H A K L A R K L W T A P E H L R E G K S M H P G G T P K G     |
| PDGFR | 829 | - - - G S T Y L L K W M A P E S I - - - - - F N S L Y T T L S     |
| c-fes | 721 | - - - G L R Q V E V K W T A P E A L - - - - - N Y G R Y S S E S   |
| c-mos | 236 | S - Y P L G G T Y T H R A P E L L - - - - - K G E G V T P K A     |
| GC    | 736 | I Y S F S I L T E M Y S - R Q E P F H E N D L E L A D I I A R     |
| PDGFR | 854 | V W S F G L L W E I F T L G G T P Y P E L - P M N D Q F Y N A     |
| c-fes | 746 | V W S F G L L W E I F T L G G T P Y P E L - L S N Q Q T R E F     |
| c-mos | 262 | I Y S F A I T L W Q M T T - K Q A P Y S - - - G E R Q H I L Y A   |
| GC    | 767 | V S K G E V P P Y R P V L N A V N E A A P D C V L T A R A C W     |
| PDGFR | 885 | I K R G Y R M A Q P A H - - - - - A S D E I Y E I N Q K C W       |
| c-fes | 776 | V E K G G R L P C P E L - - - - - C P D A V F R L Q C W           |
| c-mos | 290 | V V A Y D L R P S L S A A V F E D S L P G Q R L G D V Q R C W     |
| GC    | 799 | V E D P M E R P N I I E V R T M L A P L Q K G L K / 142           |
| PDGFR | 910 | E E K F E T R P P F S Q L V L L L E R L L G E G Y / 133           |
| c-fes | 801 | A Y E P G Q P S F S T I Y Q E L Q S I R K R H R (end)             |
| c-mos | 322 | R P S A A Q R P S A R L L L V D L T S L K A E L G (end)           |

native protein also alters its apparent size<sup>9-11</sup> (Fig. 1b). Sequences of experimentally derived *Staphylococcus aureus* V8 peptide fragments (Fig. 1 legend) are contained within the open reading frame (Fig. 3), and with the exception of peptide 1, originate after a glutamate residue, as expected for peptides generated with this protease. The peptide sequences match the cDNA-derived amino-acid sequence for the nine experimentally determined amino acid sequences except for residues 4 and 6 of peptide 2. These residues were identified to be serine and asparagine, respectively, but predicted from the complementary DNA sequence to be aspartic acid at both positions.

Proceeding toward the carboxy-terminus from the hydrophobic signal sequence, hydropathic analysis demonstrates a major hydrophobic stretch between residues 479 and 507 that may represent a membrane spanning region (data not shown). This hydrophobic sequence is immediately followed by three basic residues, Arg-Lys-Arg, a feature of the primary sequence of the cytoplasmic region of many membrane proteins. This sequence may represent the stop transfer signal anchoring the guanylate cyclase in the membrane during biosynthesis<sup>14</sup>. Hydrophobic sequences preceding the proposed transmembrane

region are not followed by the potential stop transfer signals. The 28-amino-acid potential transmembrane region, therefore, divides the protein into a 478-residue amino-terminal domain and a 459-residue carboxy-terminal domain. The specific reaction of amino-terminal synthetic peptide antiserum with guanylate cyclase of intact spermatozoa (Table 1) is consistent with the predicted extracellular exposure of the amino-terminal domain of guanylate cyclase. The proposed intracellular carboxyl-terminal domain of the protein has several regions of hydrophobic amino acids, and a high abundance (approximately 9%) and clustering of serine residues (Fig. 3). Serine residues occur in four regions after the proposed transmembrane domain with only 8 out of 41 serine residues occurring at more than three amino acids from another serine. The 95 amino acids of the carboxy terminus are composed of 21% serine (Fig. 3).

There is no extensive sequence similarity of sea-urchin sperm guanylate cyclase with adenyl cyclase from yeast (*Saccharomyces cerevisiae*)<sup>16</sup> or bacteria (*Escherichia coli*)<sup>17</sup>, or with other proteins. However, there are distinct sequence similarities between sea-urchin spermatozoan guanylate cyclase and highly conserved regions of protein kinases<sup>18-20</sup>, illustrated by a com-

parison of sea-urchin guanylate cyclase with the mouse platelet-derived growth factor (PDGF) receptor<sup>21</sup> and human c-fes<sup>22</sup> tyrosine kinases, and human c-mos<sup>23</sup> serine/threonine protein kinase amino-acid sequences (Fig. 4). Residues 530–823 of guanylate cyclase are from approximately 23% (PDGF receptor) to 25% (c-fes and c-mos) identical with the catalytic domains of these protein kinases; the greatest similarities are with members of the protein tyrosine kinase and mos/raf protein families (approximately 19% to 25%) as compared to other members of the serine protein kinase families (approximately 14% to 21%). The sequence identities are not distributed randomly; of 33 residues highly conserved in all reported protein kinases<sup>24</sup> (Fig. 4), 26 are retained in guanylate cyclase. Replacements in the guanylate cyclase sequence compared to the highly conserved protein kinase residues are as follows: Ser is substituted for Gly at three positions (541, 544 and 740), Gln is substituted for hydrophobic residues at two positions (residues 538 and 560), Arg is substituted for Ser/Thr/Pro (position 711), and His replaces Asp (residue 669); these substitutions may be important in defining the different catalytic functions of guanylate cyclase and protein kinases since these positions are invariant or conserved in all published protein kinase sequences<sup>24</sup>. The sequence similarity of guanylate cyclase with protein kinases places guanylate cyclase in the protein kinase family, and may imply a similar architecture of their respective nucleotide binding sites. The results presented here also raise a question concerning the reliability of classifying a protein as a protein kinase based on sequence similarity alone. Conversely, it will be important to determine whether or not guanylate cyclase has phosphotransferase activity.

We thank Peter Ng and Parkash Jhurani for synthesizing oligonucleotide probes. Amrik Khatri, Jan Jordan, Nancy C.

Saul and D. Janette Tubb for technical assistance, Jeanne Arch and Penny Stelling for typing the manuscript and Donna Dangott and Carol Morita for artwork. We thank Steven Hanks, Anne Marie Quinn and Tony Hunter for communicating results prior to publication. This work was supported in part by the NIH.

Received 22 April; accepted 4 July 1988.

- Shimomura, H., Dangott, L. J. & Garbers, D. L. *J. biol. Chem.* **261**, 15778–15782 (1986).
- Kuno, T. *et al. J. biol. Chem.* **261**, 5817–5823 (1986).
- Paul, A. K., Marala, R. B., Jaiswal, R. K. & Sharma, R. K. *Science* **235**, 1224–1226 (1987).
- Takayanagi, R. *et al. J. biol. Chem.* **262**, 12104–12113 (1986).
- Bentley, J. K., Shimomura, H. & Garbers, D. L. *Cell* **45**, 281–288 (1986).
- Bentley, J. K., Tubb, D. J. & Garbers, D. L. *J. biol. Chem.* **261**, 14859–14862 (1986).
- Radany, E. W., Gerzer, R. & Garbers, D. L. *J. biol. Chem.* **253**, 8346–8351 (1983).
- IUPAC-IUB Committee on Biochemical Nomenclature, *Biochem. J.* **113**, 1 (1969).
- Ward, G. E., Garbers, D. L. & Vacquier, V. D. *Science* **227**, 768–770 (1985).
- Suzuki, N. *et al. J. biol. Chem.* **259**, 14874–14879 (1984).
- Ward, G. E. & Vacquier, V. D. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5578–5582 (1983).
- Kozak, M. *Nucleic. Acids Res.* **9**, 5233–5262 (1981).
- Watson, M. E. E. *Nucleic Acids Res.* **12**, 5145–5164 (1984).
- Sabatini, D. D., Kreibich, G., Morimoto, T. & Adesnik, M. *J. Cell Biol.* **92**, 1–22 (1982).
- Ramarao, C. S. & Garbers, D. L. *J. biol. Chem.* **263**, 1524–1529 (1988).
- Kataoka, T., Broek, D. & Wigler, M. *Cell* **43**, 493–505 (1985).
- Aiba, H., Mori, K., Tanaka, M., Ooi, T., Roy, A. & Danchin, A. *Nucleic Acids Res.* **12**, 9427–9440 (1984).
- Hunter, T. & Cooper, J. A. in *The Enzymes* (eds Boyer, P. D. & Krebs, E. G.) **17**, 191–246 (Academic, New York, 1986).
- Hunter, T. *Cell* **50**, 823–829 (1987).
- Knopf, J. L. *et al. Cell* **46**, 491–502 (1986).
- Yarden, Y. *et al. Nature* **323**, 226–232 (1986).
- Roebroek, A. J. M. *et al. EMBO J.* **4**, 2897–2903 (1985).
- Watson, R., Oskarsson, M. & Van Woude, G. F. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4078–3082 (1982).
- Hanks, S. K., Quinn, A. M. & Hunter, T. *Science* **241**, 42–52 (1988).
- Garbers, D. L. & Murad, F. *Adv. Cyclic Nucleotide Res.* **10**, 57–67 (1979).
- McLean, J. W. *et al. Nature* **330**, 132–137 (1987).
- Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161–178 (1980).

## A mutation that prevents GTP-dependent activation of the $\alpha$ chain of $G_s$

R. Tyler Miller, Susan B. Masters, Kathleen A. Sullivan, Barry Beiderman & Henry R. Bourne

Departments of Pharmacology, Medicine, and the Cardiovascular Research Institute, University of California, San Francisco, California 94143-0450, USA

Membrane-bound G proteins carry information from receptors on the outside of cells to effector proteins inside cells. The  $\alpha$  subunits of these heterotrimeric proteins bind and hydrolyse GTP and control the specificity of interactions with receptor and effector elements<sup>1,2</sup>. Signalling by G proteins involves a cycle in which the inactive  $\alpha\beta\gamma$ -GDP complex dissociates to produce  $\alpha^*$ -GTP, which is capable of activating the effector enzyme or ion channel; the  $\alpha^*$ -GTP complex hydrolyses bound GTP and reassociates with  $\beta\gamma$  to form the inactive complex. We have characterized a mutation that interrupts this GTP-driven cycle in  $\alpha_s$ , the  $\alpha$ -chain of  $G_s$ , the G protein that stimulates adenylyl cyclase. The mutation converts a glycine to an alanine residue in the presumed GDP-binding domain of  $\alpha_s$ . The location and biochemical consequences of this mutation suggest a common mechanism by which binding of GTP or ATP may induce changes in the conformation of a number of nucleoside triphosphate binding proteins.

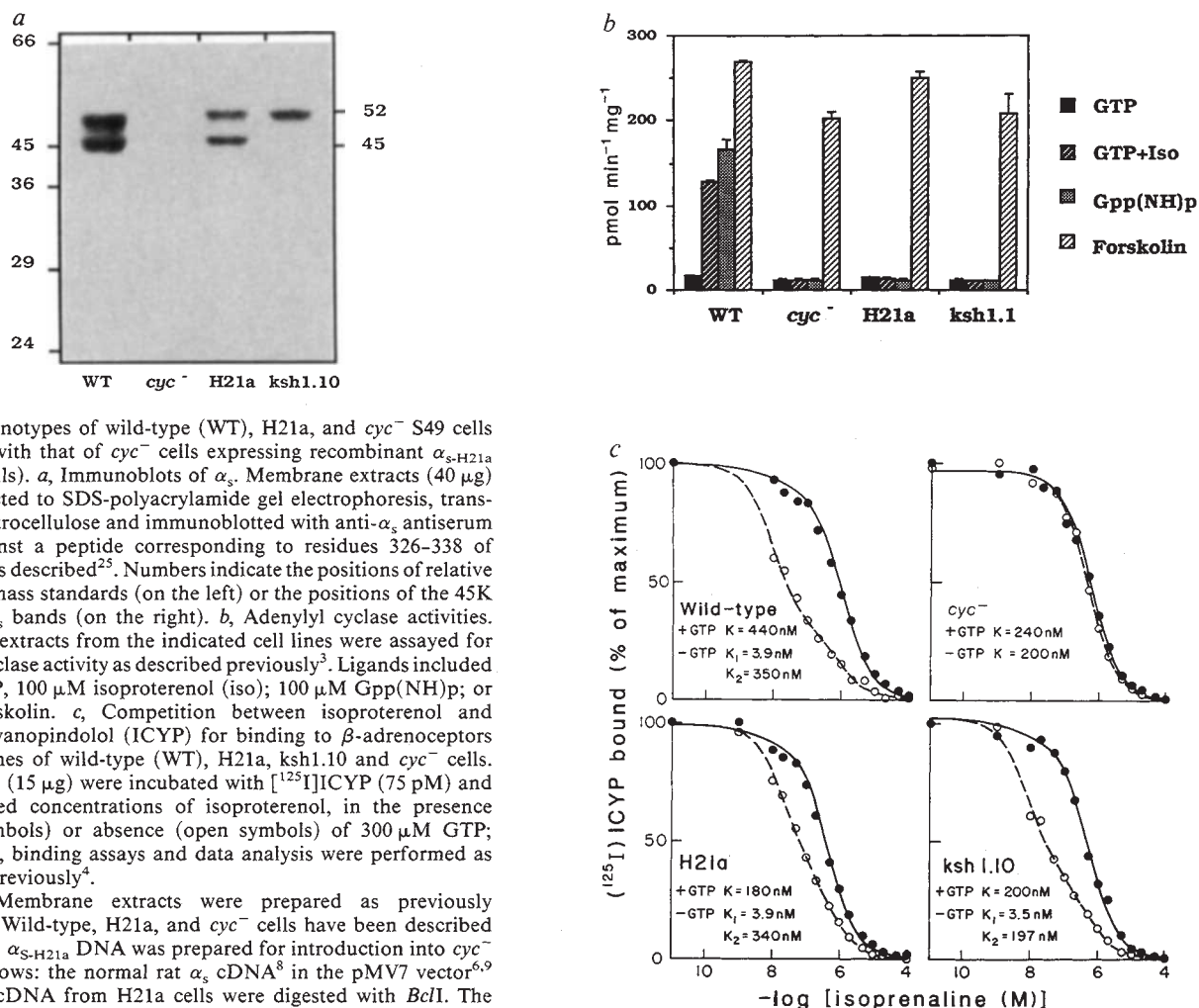
This mutation is responsible for a mutant phenotype, called H21a, in S49 mouse lymphoma cells. Although  $G_s$  in H21a membranes interacts normally with receptors (it can promote high-affinity binding of  $\beta$ -adrenoceptor agonists), adenylyl cyclase in H21a cells fails to respond to stimulation by agents that act through receptors or  $G_s$ , including hormonal agonists, cholera toxin, GTP analogues, and  $AlF_4^-$  ions<sup>3,4,5</sup>. Somatic

genetic analysis showed that the H21a mutation is in the gene that encodes the  $\alpha$ -chain of  $G_s$  ( $\alpha_s$ )<sup>5</sup>.

We prepared complementary DNA (cDNA) libraries from S49 wild-type and H21a mutant cells and isolated multiple partial cDNAs encoding  $\alpha_s$  from each, as described previously<sup>6</sup>. Two independent wild-type and two independent H21a cDNA inserts, each encoding approximately 70% of the open reading frame, differed in one base only: substitution of cytosine for guanosine at position 696 which would replace Gly 226 by alanine<sup>7</sup>. To rule out the possibility that an additional mutation in the unsequenced 5' part of the coding region was responsible for the H21a phenotype, we constructed a recombinant cDNA encoding the form of  $\alpha_s$  of relative molecular mass 52,000 (52K) with the H21a base substitution in codon 226 (see Fig. 1). We expressed the mutant cDNA and its wild-type counterpart in genetically  $\alpha_s$ -deficient S49 *cyc*<sup>-</sup> cells, exactly as described previously<sup>6,8,9</sup>.

Expression of the  $\alpha_{s-H21a}$  cDNA in *cyc*<sup>-</sup> produced a 52K  $\alpha_s$  polypeptide detectable by immunoblotting (Fig. 1a) or by cholera toxin-catalyzed radiolabelling with [<sup>32</sup>P]NAD<sup>+</sup> (not shown). As in the original H21a mutant, adenylyl cyclase in membranes from *cyc*<sup>-</sup> cells expressing the recombinant  $\alpha_{s-H21a}$  polypeptide were not stimulated by either GTP plus the  $\beta$ -adrenoceptor agonist, isoproterenol, or by a GTP analogue, guanylyl-5'-imidodiphosphate (Gpp[NH]p); forskolin, which acts directly on catalytic adenylyl cyclase<sup>10</sup>, did stimulate cyclic AMP (cAMP) synthesis in H21a membranes, as it does in *cyc*<sup>-</sup> (Fig. 1b). The recombinant  $\alpha_{s-H21a}$  polypeptide interacted normally with  $\beta$ -adrenoceptors, as is the case in H21a mutant cells. This interaction was demonstrated (Fig. 1c) by the high affinity of these receptors for binding agonist in the absence of guanine nucleotide and the shift to low affinity in the presence of guanine nucleotide; receptors in *cyc*<sup>-</sup> membranes bind agonist with low affinity, regardless of the presence of guanine nucleotide, because these cells lack  $\alpha_s$  altogether. These results show that





**Fig. 1** Phenotypes of wild-type (WT), H21a, and *cyc*<sup>-</sup> S49 cells compared with that of *cyc*<sup>-</sup> cells expressing recombinant  $\alpha_{s-H21a}$  (ksh1.10 cells). *a*, Immunoblots of  $\alpha_s$ . Membrane extracts (40  $\mu$ g) were subjected to SDS-polyacrylamide gel electrophoresis, transferred to nitrocellulose and immunoblotted with anti- $\alpha_s$  antiserum raised against a peptide corresponding to residues 326–338 of mouse  $\alpha_s$ , as described<sup>25</sup>. Numbers indicate the positions of relative molecular mass standards (on the left) or the positions of the 45K and 52K  $\alpha_s$  bands (on the right). *b*, Adenylyl cyclase activities. Membrane extracts from the indicated cell lines were assayed for adenylyl cyclase activity as described previously<sup>3</sup>. Ligands included 50  $\mu$ M GTP, 100  $\mu$ M isoproterenol (iso); 100  $\mu$ M Gpp(NH)p; or 10  $\mu$ M forskolin. *c*, Competition between isoproterenol and [<sup>125</sup>I]iodocyanopindolol (ICYP) for binding to  $\beta$ -adrenoceptors in membranes of wild-type (WT), H21a, ksh1.10 and *cyc*<sup>-</sup> cells. Membranes (15  $\mu$ g) were incubated with [<sup>125</sup>I]ICYP (75 pM) and the indicated concentrations of isoproterenol, in the presence (closed symbols) or absence (open symbols) of 300  $\mu$ M GTP; incubations, binding assays and data analysis were performed as described previously<sup>4</sup>.

**Methods.** Membrane extracts were prepared as previously described<sup>3</sup>. Wild-type, H21a, and *cyc*<sup>-</sup> cells have been described previously<sup>5</sup>.  $\alpha_{s-H21a}$  DNA was prepared for introduction into *cyc*<sup>-</sup> cells as follows: the normal rat  $\alpha_s$  cDNA<sup>8</sup> in the pMV7 vector<sup>6,9</sup> and an  $\alpha_s$  cDNA from H21a cells were digested with *Bcl*I. The 8,250 base pairs (bp) vector and rat  $\alpha_s$  fragment (without the *Bcl*I–*Bcl*I fragment) and the 727 bp *Bcl*I–*Bcl*I fragment from the H21a  $\alpha_s$  cDNA were purified and ligated. The deduced amino acid sequence of the chimera was identical to that of a previously sequenced S49  $\alpha_s$  cDNA<sup>7</sup> except for the replacement of glycine by alanine at residue 226. The  $\alpha_{s-H21a}$  DNA was expressed in *cyc*<sup>-</sup> cells as described<sup>6</sup>. The *cyc*<sup>-</sup> cell line used in these experiments was a subclone of the *cyc*<sup>-</sup> clone used previously.

replacement of Gly 226 in  $\alpha_s$  by alanine is responsible for the H21a phenotype.

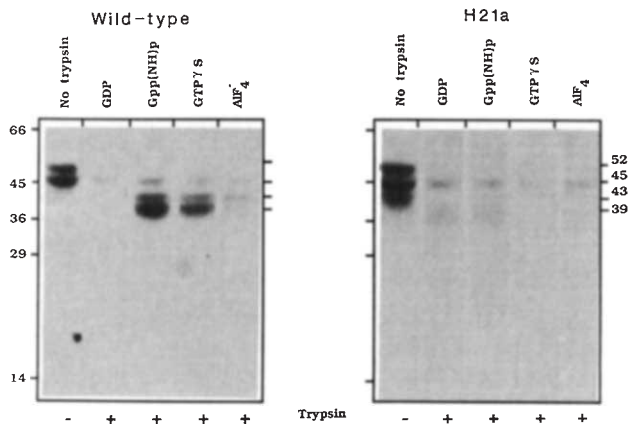
We expected that the H21a mutation would lie in a region of  $\alpha_s$  that interacts with adenylyl cyclase. The region surrounding Gly 226 is unlikely to be involved in such an interaction however, because its sequence is almost identical in G protein  $\alpha$ -chains that interact with different effector molecules<sup>1,2</sup>. This conservation of primary structure suggested instead that the region surrounding Gly 226 might be involved in the conformational change that occurs when G proteins are activated by binding GTP. This conformational change has been documented in two ways. GTP was shown to induce a Mg<sup>2+</sup>-dependent increase in intrinsic fluorescence of the  $\alpha$ -chain of G<sub>o</sub> (refs 11, 12). Resistance of G protein  $\alpha$ -chains to proteolysis provided a second kind of evidence: in their GTP-bound active states, the  $\alpha$ -chains of transducin<sup>13,14</sup> and G<sub>o</sub> (ref. 15) resist cleavage by trypsin at a site that corresponds to Arg-232 of  $\alpha_s$ ; in each case this site is cleaved by trypsin in the inactive GDP-bound conformation of the G protein.

We investigated conformation by assessing sensitivity of the  $\alpha_s$  chain to tryptic cleavage (Fig. 2). In the presence of Mg<sup>2+</sup> and non-hydrolysable GTP analogues, trypsin trimmed the 52K and 45K forms of wild type  $\alpha_s$  to 41K and 39K, respectively, and these proteolytic products were protected from further degradation; AlF<sub>4</sub><sup>-</sup> also protected, although less efficiently. Consistent with the Mg<sup>2+</sup> requirement for GTP-induced conforma-

tional change in  $\alpha_o$  (refs 11, 12), protection of  $\alpha_s$  by these agents also required Mg<sup>2+</sup> (not shown). In the presence of GDP, trypsin degraded both the 52K and 45K  $\alpha_s$  polypeptides in wild type membranes. In contrast, when H21a membranes were exposed to trypsin in the presence of GTP analogues or AlF<sub>4</sub><sup>-</sup>, no protected polypeptide bands of 41K or 39K were seen (Fig. 2).

The most likely interpretation of the trypsin experiments is that  $\alpha_{s-H21a}$ —unlike normal  $\alpha_s$ ,  $\alpha_i$  or  $\alpha_o$ —cannot assume the active, GTP-bound conformation that protects a large fragment of the polypeptide from cleavage by trypsin. A separate observation rules out the possibility that this defect of  $\alpha_{s-H21a}$  results from failure to bind GTP or GTP analogues: like GTP itself (Fig. 1c), Gpp(NH)p abolishes high affinity binding of  $\beta$ -adrenoceptor agonists to receptors in H21a membranes (not shown), indicating that  $\alpha_{s-H21a}$  does bind the GTP analogue.

GTP-induced activation of G protein  $\alpha$ -chains is accompanied by decreased affinity for binding  $\beta\gamma$ <sup>1,2</sup>. We therefore assessed dissociation of  $\alpha_s$  from  $\beta\gamma$  by examining the sedimentation rate of  $\alpha$ -chains in sucrose gradients in the presence or absence of activating ligands (Fig. 3). Wild-type  $\alpha_s$  sedimented more rapidly (4.6  $\pm$  0.2 S; standard error of the mean (s.e.m.), *n* = 3) in the absence of AlF<sub>4</sub><sup>-</sup> ion than in its presence (2.7  $\pm$  0.5 S; s.e.m., *n* = 3); this reduction in sedimentation rate is thought to reflect AlF<sub>4</sub><sup>-</sup> induced dissociation of  $\alpha_s$  from  $\beta\gamma$ <sup>16</sup>. In contrast, AlF<sub>4</sub><sup>-</sup> caused no change in the sedimentation of  $\alpha_{s-H21a}$  (4.4  $\pm$



**Fig. 2** Immunoblot showing tryptic cleavage patterns of  $\alpha_s$  in wild type and H21a membranes. Where indicated, concentrations were 50  $\mu$ M GDP, 50  $\mu$ M Gpp(NH)p, 50  $\mu$ M GTP $\gamma$ S, 100  $\mu$ M isoproterenol and AIF $_4$  (10 mM NaF and 10  $\mu$ M AlCl $_3$ ). Numbers at left margin indicate migration of relative molecular mass standards and numbers at right margin indicate relative molecular masses of tryptic fragments.

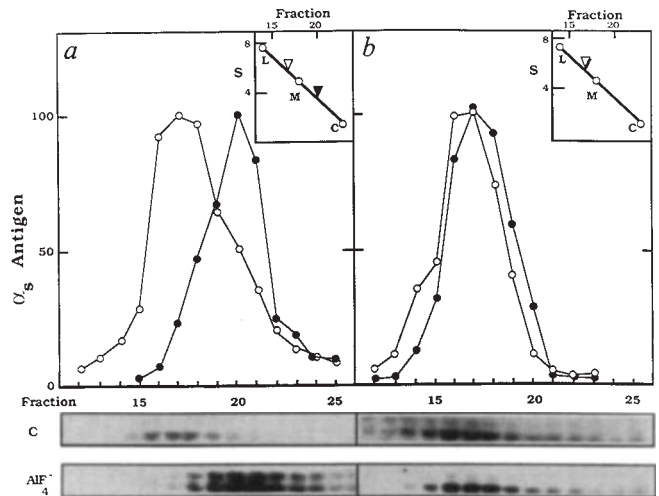
**Methods.** Membrane extracts of wild-type and H21a cells were prepared as previously described<sup>3</sup>. Membranes were diluted to a concentration of 2.5 mg ml $^{-1}$  in buffer containing 20 mM HEPES (pH 8, 4.0 mM MgCl $_2$ , 1.0 mM EDTA and lubrol PX 0.08% (w/v). Guanine nucleotides, analogues, AIF $_4$ , and isoproterenol were added to the final concentrations given above. The membranes were incubated for 30 min at 30 °C. Trypsin was added to a final concentration of 200  $\mu$ g ml $^{-1}$  and the mixture incubated for 5 min at 30 °C. The samples, 40  $\mu$ g of membrane protein per lane, were run on a 10% SDS-polyacrylamide gel, and immunoblotted as described (see Fig. 3)<sup>25</sup>.

0.2 S; s.e.m.,  $n=3$  in both conditions). Similar results (not shown) were obtained with Gpp(NH)p instead of AIF $_4$ .

Our studies show that the H21a mutation prevents a GTP-induced conformational change that is responsible for dissociation of the  $\alpha$ -chain from  $\beta\gamma$  as well as for resistance to trypsin and activation of adenyl cyclase. This change is distinct from the guanine nucleotide-induced alteration in conformation that causes  $\alpha\beta\gamma$  to dissociate from receptors, as  $G_s$  in H21a can couple reversibly to receptors. These results also suggest that the activation of G proteins by guanine nucleotides is sequential, with dissociation of the GTP-bound  $\alpha\beta\gamma$  heterotrimer from receptor preceding separation of  $\alpha^*$ -GTP from  $\beta\gamma$  and subsequent activation of the effector.

Figure 4 shows a model<sup>17</sup> of the three-dimensional structure of the GDP-binding domain of a G protein  $\alpha$ -chain, based on the crystal structure of the corresponding domain of bacterial EF-Tu<sup>18,19</sup>. The proposed structure also resembles the recently solved crystal structure<sup>20</sup> of a second GTP-binding protein, p21<sup>Ha-ras</sup>. In the model, GDP nestles into a binding site formed by loops between  $\beta$ -strands and  $\alpha$ -helices. Similarities in the amino acid sequences in each of the loops suggest that corresponding residues in the G protein  $\alpha$ -chains contact the GDP molecule in a similar way to that observed in EF-Tu and p21<sup>ras</sup>.

We propose that the ability of  $\alpha_s$  to discriminate between GTP and GDP depends upon a GTP-induced movement of the region (residues 222–228) shaded in Fig. 4 and that this movement initiates a conformational change that is transmitted to the rest of the protein. The methyl side chain inserted by the H21a mutation at a turn in the centre of this region presumably blunts or deflects the GTP-induced movement and thereby blocks conversion of the protein into its active form.

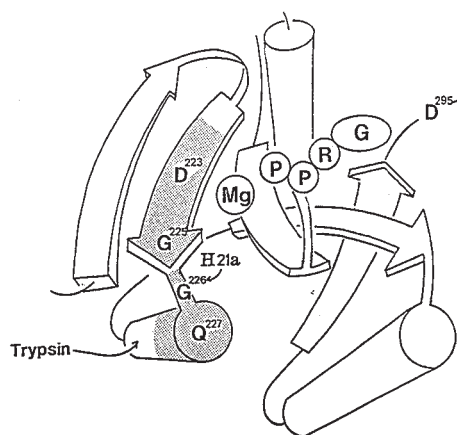


**Fig. 3** Sucrose gradient sedimentation profiles of lubrol extracts from wild type (a) and H21a (b) membranes treated with (closed symbols) or without (open symbols) AIF $_4$  (10 mM NaF, 10  $\mu$ M AlCl $_3$ ). Lubrol extracts of the S49 cell membranes were applied to sucrose gradients with enzyme standards. The fractions were subjected to electrophoresis in SDS gels and immunoblotted with anti- $\alpha_s$  antiserum, as described<sup>25</sup>. Autoradiograms of the immunoblots were subjected to densitometry (LKB Ultrosan XL). Areas under the optical density curves representing the 45K bands for each fraction were compared by expressing each value as a percentage of the maximal value detected on the autoradiogram. Insets (top right in each figure) show the migration of enzyme standards (cytochrome c [C], 1.9S; malate dehydrogenase [M], 4.3 S; lactate dehydrogenase [L], 7.45S. The arrows indicate peaks of  $\alpha_s$  in the gradients, as determined by densitometry of immunoblots using anti- $\alpha_s$  antiserum. The relevant portions of the immunoblots from gradient fractions without (C) or with AIF $_4$  are shown below the graphs.

**Methods.** Membrane extracts of wild type and H21a S49 cells were prepared as previously described<sup>3</sup>. Membrane protein (1 mg) was extracted for 2 h at 4 °C in a volume of 500  $\mu$ l with 0.4% Lubrol (w/v) in a buffer containing 100 mM NaCl; 10 mM MgCl $_2$ ; 1.25 mM EDTA; 20 mM HEPES (pH 8.0); 2.0 mM  $\beta$ -mercaptoethanol; 1.0 mM phenyl methylsulphonate; 50  $\mu$ g ml $^{-1}$  leupeptin; and 50  $\mu$ g ml $^{-1}$  soybean trypsin inhibitor. The membranes were centrifuged in an airfuge (Beckman) for 10 min at 20 psi. AIF $_4$  or vehicle was added to 200  $\mu$ l aliquots of the membrane extract supernatant, and the mixture incubated for 10 min at room temperature. Membrane extracts were combined with 35  $\mu$ l aliquots of the marker enzymes and loaded on to 5–20% sucrose gradients made with the buffer described above, except that Lubrol was 0.1% (w/v). Gradients for the activated wild-type and H21a membranes also contained 10 mM NaF and 10  $\mu$ M AlCl $_3$ . Gradients were centrifuged for 15 h at 40,000 r.p.m. at 4.0 °C in a Beckman SW55ti rotor in an L8 M ultracentrifuge. Fractions (200  $\mu$ l) were collected. After removal of 35  $\mu$ l aliquots for marker enzyme assays, the fractions were precipitated with trichloroacetic acid, resuspended in base and subjected to gel electrophoresis and immunoblotting, as described<sup>25</sup>. Marker enzyme assays were performed as previously described<sup>5</sup>.

The highly conserved nature of amino acid sequence<sup>1,2,21–23</sup> in this region extends to all the G protein  $\alpha$ -chains, as well as adenylate kinase, EF-Tu, and—to a lesser extent—p21<sup>ras</sup>. This conservation suggests that the region may play a common mechanical role in several GTP- and ATP-binding proteins. It is plausible, for example, that the conserved aspartate residue (corresponding to Asp-223 in  $\alpha_s$ , located in the  $\beta$ -strand upstream from the mutated residue in H21a) forms a contact between this region and the bound nucleoside triphosphate or diphosphate. The crystal structure of EF-Tu<sup>18,19</sup> reveals a Mg $^{2+}$





**Fig. 4** Schematic model of the proposed guanine nucleotide-binding region of G protein  $\alpha$ -chains. GDP is shown with its guanine ring (G) interacting with Asp 295 and its phosphoryls (P) interacting, by a  $Mg^{2+}$  salt bridge, with Asp 223; R = ribose portion of GDP. The shaded structure ( $\beta$ -strand  $\rightarrow \alpha$ -helix) is the region whose movement is thought to participate in the GTP-induced conformational change of  $\alpha$ . H21a indicates the amino acid residue (Gly-226) that is replaced by alanine in the mutant  $\alpha_s$ . G-225 (Gly 225) and Q-227 (Gln 227) are equivalent, respectively, to Ala 59 and Gln 61 of p21<sup>ras</sup> positions at which amino acid substitutions activate the protein<sup>24,26</sup>. 'Trypsin' indicates the presumed conformation-sensitive trypsin cleavage site, based on results with  $\alpha_i$  and  $\alpha_o$ <sup>13-15</sup>.

ion that forms a salt bridge between the phosphoryls of GDP and this aspartate residue (see Fig. 4). This  $Mg^{2+}$  ion may serve as the 'button' used by the  $\gamma$ -phosphoryl of GTP to exert pressure on the region shaded in Fig. 4. This proposal is consistent with the absolute requirement for  $Mg^{2+}$  ions for the GTP-induced change in  $\alpha$ -chain conformation detected by tryptophan fluorescence<sup>11,12</sup> or by resistance to tryptic cleavage.

Our results identify a conserved region of GTP-binding proteins whose change in conformation lies at the heart of GTP-dependent transmembrane signalling. More precise definition of this key conformational change will require additional information from biochemical and biophysical experiments.

Received March 28; accepted 12 July 1988.

- Gilman, A. G. *A. Rev. Biochem.* **56**, 615-649 (1987).
- Stryer, L. & Bourne, H. R. *A. Rev. Cell Biol.* **2**, 391-419 (1986).
- Salomon, M. R. & Bourne, H. R. *Molec. Pharmac.* **19**, 109-116 (1981).
- Bourne, H. R., Kaslow, D., Kaslow, H. R., Salomon, M. R. & Licko, V. *Molec. Pharmac.* **20**, 435-441 (1981).
- Bourne, H. R., Beiderman, B., Steinberg, F. & Brothers, V. M. *Molec. Pharmac.* **22**, 204-210 (1982).
- Sullivan, K. A. *et al. Nature* **330**, 758-760 (1987).
- Rall, T. & Harris, B. A. *FEBS Lett.* **224**, 365-371 (1987).
- Jones, D. T. & Reed, R. R. *J. Biol. Chem.* **262**, 14241-14249 (1987).
- Kirschmeier, P. T., Housey, G. M., Johnson, M. D., Perkins, A. S. & Weinstein, I. B. *DNA* **7**, 219-225 (1988).
- Seamon, K. & Daly, J. W. *J. Biol. Chem.* **257**, 9799-9801 (1981).
- Higashijima, T. *et al. J. Biol. Chem.* **262**, 752-756 (1987).
- Higashijima, T., Ferguson, K. M., Smigel, M. D. & Gilman, A. G. *J. Biol. Chem.* **262**, 757-761 (1987).
- Fung, B. K.-K. & Nash, C. R. *J. Biol. Chem.* **258**, 10503-10510 (1983).
- Bigay, J., Deterre, P., Pfister, C. & Chabre, M. *EMBO J.* **6**, 2907-2913 (1987).
- Hurley, J. B., Simon, M. I., Teplow, D. B., Robishaw, J. D. & Gilman, A. G. *Science* **226**, 860-862 (1984).
- Howlett, A. C. & Gilman, A. G. *J. Biol. Chem.* **255**, 2861-2866 (1980).
- Masters, S. B., Stroud, R. M. & Bourne, H. R. *Protein Engineering* **1**, 47-54 (1986).
- Jurnak, F. *Science* **230**, 32-36 (1985).
- La Cour, T. F. M., Nyborg, J., Thrup, S. & Clark, B. F. *EMBO J.* **4**, 2385-2388 (1985).
- De Vos, A. M. *et al. Science* **239**, 888-893 (1988).
- Fry, D. C., Kuby, S. A. & Mildvan, A. S. *Proc. natn. Acad. Sci. U.S.A.* **83**, 907-911 (1986).
- Dietzel, C. & Kurjan, J. *Cell* **50**, 1001-1010 (1987).
- Miyajima, I. *et al. Cell* **50**, 1011-1019 (1987).
- McCormick, F. *et al. Science* **230**, 78-82 (1985).
- Masters, S. B. *et al. Science* **241**, 448-451 (1988).
- Der, C. J., Finkel, T. & Cooper, G. M. *Cell* **44**, 167-176 (1986).

## Insulin-stimulated MAP-2 kinase phosphorylates and activates ribosomal protein S6 kinase II

Thomas W. Sturgill\*, L. Bryan Ray\*, Eleanor Erikson† & James L. Maller†‡

\* Departments of Internal Medicine and Pharmacology, University of Virginia School of Medicine, Charlottesville, Virginia 22903, USA

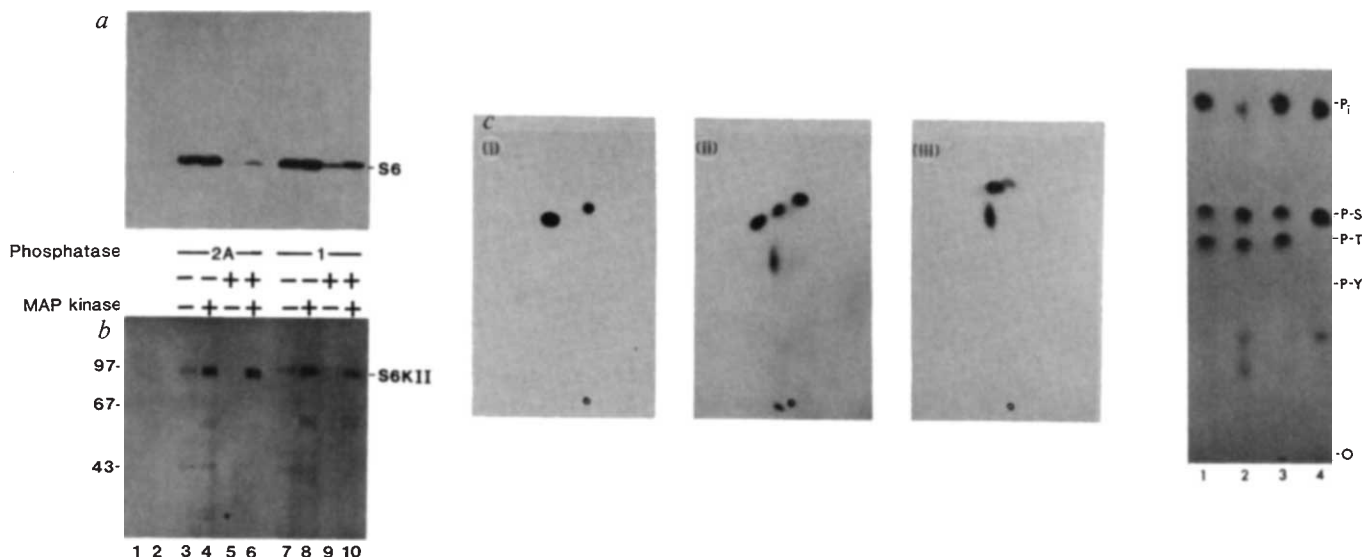
† Department of Pharmacology, University of Colorado School of Medicine, Denver, Colorado 80262, USA

Ribosomal protein S6 is a component of the eukaryotic 40S ribosomal subunit that becomes phosphorylated on multiple serine residues in response to a variety of mitogens, including insulin, growth factors, and transforming proteins of many oncogenic viruses. Recently, an activated S6 kinase (S6 K II) has been purified to homogeneity from *Xenopus* eggs<sup>1</sup>, and characterized immunologically<sup>2</sup> and at the molecular level<sup>3</sup>. Purified S6 K II can be deactivated *in vitro* by incubation with either protein phosphatase 1 or protein phosphatase 2A. Reactivation and phosphorylation of S6 K II occurs *in vitro* with an insulin-stimulated microtubule-associated protein-2 (MAP-2) protein kinase which is itself a phosphoprotein that can be deactivated by protein phosphatase 2A. These studies suggest that a step in insulin signalling involves sequential activation by phosphorylation of at least two serine/threonine protein kinases.

In all cases so far, extracts of mitogen-stimulated cells have been shown to contain elevated S6 kinase activity<sup>4-7</sup>. A major question at present concerns how S6 kinases are activated, particularly the very rapid increases elicited by insulin in oocytes and other cells<sup>6-9</sup>. Since the insulin receptor kinase has an absolute specificity for tyrosine residues, there must exist a pathway by which an activated tyrosine protein kinase can activate a serine kinase. Several mechanisms could account for the activation of S6 kinases by insulin. Direct activation of the S6 kinase by phosphorylation on tyrosine residues is an attractive possibility, but purified S6 K II is not a substrate *in vitro* for insulin receptor kinase, pp60<sup>src</sup>, or the tyrosine kinase associated with Abelson murine leukaemia virus<sup>10</sup>. Other evidence against a direct activation by tyrosine phosphorylation of S6 K II is the absence of phosphotyrosine in immunoprecipitates of the enzyme from <sup>32</sup>P-labelled, insulin-treated oocytes<sup>10</sup>.

Clear support for an indirect mechanism comes from the finding that the serine/threonine-specific protein phosphatases 1 and 2A can deactivate S6 K II *in vitro*<sup>10</sup>. Recently, it was reported that S6 kinase from 3T3 cells could also be deactivated by protein phosphatase 2A<sup>11</sup>, indicating that S6 kinases in general are regulated by direct phosphorylation rather than by second messengers. Although many protein kinases can be activated by autophosphorylation, this is not the case for purified S6 K II *in vitro*<sup>1</sup>. This finding raised the possibility that a serine/threonine kinase may be responsible for the activation of S6 K II. A number of protein-serine/threonine kinases, including cyclic AMP-dependent protein kinase, and protein kinase C were unable to reactivate S6 K II *in vitro* after phosphatase treatment (data not shown), suggesting the activating serine/threonine protein kinase might be a relatively specific enzyme under acute regulation by insulin. The existence of such an enzyme in insulin-treated cells was reported by Ray and Sturgill, who identified in 3T3-L1 cells an insulin-stimulated kinase able to phosphorylate microtubule-associated protein-2 (MAP-2) *in vitro* on serine and threonine residues<sup>12</sup>. This kinase is activated transiently by insulin before S6 kinase activation and is unable to phosphorylate histones, casein, S6, or any of a number of substrates whose phosphorylation is increased by insulin<sup>13</sup>.

‡ To whom correspondence should be addressed.



**Fig. 1** Deactivation of S6 kinase by phosphatase treatment and reactivation and phosphorylation by MAP-2 kinase. *a*, S6 kinase activity. S6 K II was incubated with either phosphatase 2A or phosphatase 1 as indicated. Controls (— phosphatase in figure) were treated with phosphatase 2A in the presence of NaF, or phosphatase 1 in the presence of inhibitor-2. After termination of the phosphatase reaction, the samples were incubated with ATP without or with MAP kinase, and then with 40S subunits and [ $\gamma$ - $^{32}$ P]ATP. Phosphorylation of S6 was quantified as described previously<sup>1</sup>. Lanes 1 and 2 are 40S subunits and [ $\gamma$ - $^{32}$ P]ATP without and with MAP kinase, respectively. The c.p.m. in S6 were: lane 1, 97; lane 2, 181; lane 3, 2,448; lane 4, 3,692; lane 5, 226; lane 6, 797; lane 7, 3,413; lane 8, 4,304; lane 9, 1,043; lane 10, 1,983. *b*, Phosphorylation of S6 K II. Samples corresponding to those in *a* were treated with phosphatase and then phosphorylated in the presence of [ $\gamma$ - $^{32}$ P]ATP without (odd-numbered lanes) or with (even-numbered lanes) MAP kinase as indicated. *c*, Phosphopeptide mapping. S6 K II was subjected to two-dimensional phosphopeptide mapping as described below after autophosphorylation (i), after phosphorylation by MAP kinase without phosphatase 1 pretreatment (ii) and after phosphorylation by MAP kinase after phosphatase 1-mediated deactivation (iii). Fractionation of a mixture of samples (i) and (iii) yielded a pattern similar to that shown in (ii) (data not shown). Longer exposure of the map shown in (iii) revealed the presence of phosphopeptides corresponding to those in (i). The dashed circle denotes the origin. *d*, Phosphoamino acid analysis. S6 K II prepared as described below was subjected to 1-dimensional phosphoamino acid analysis at pH 3.5. Lane 1, S6 K II phosphorylated by MAP kinase after phosphatase 2A-mediated deactivation; lane 2, S6 K II phosphorylated by MAP kinase without phosphatase treatment; lane 3, S6 K II phosphorylated by MAP kinase after deactivation by phosphatase 1; lane 4, S6 K II incubated with [ $\gamma$ - $^{32}$ P]ATP alone (autophosphorylation).

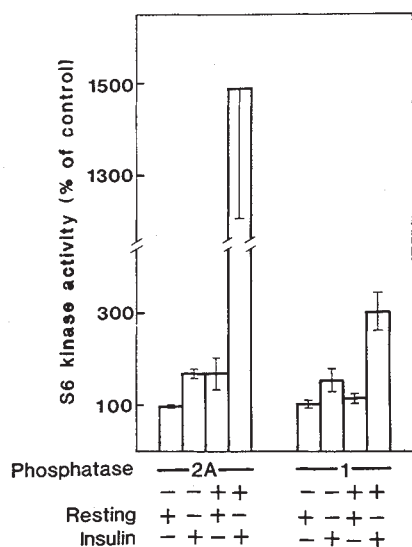
**Methods.** The preparation of S6 K II, MAP kinase, the catalytic subunits of phosphatases 1 and 2A, and inhibitor 2 were as described previously<sup>1,13,21,22</sup>. Phosphatase treatments were performed in a final volume of 100  $\mu$ l containing 50 mM Tris-HCl (pH 7.0), 30 mM 2-mercaptoethanol, 0.01% Brij-35, 0.5 unit of either protein phosphatase 2A or protein phosphatase 1 and 1 unit of S6 K II. After incubation at 30°C for 15 min phosphatase activity was inhibited by the addition of NaF to 10 mM or 240 units of inhibitor 2. One unit of either phosphatase releases 1 nmol min<sup>-1</sup> of  $^{32}$ P<sub>i</sub> from phosphorylase *a* under standard conditions<sup>21</sup>, and one unit of S6 K II transfers 1 pmol of phosphate min<sup>-1</sup> into S6 under standard conditions<sup>1</sup>. Control samples (— phosphatase in figure) contained phosphatase treated with NaF or inhibitor 2 throughout the incubation period. Under these conditions the activity of both phosphatases against phosphorylase *a* was inhibited by more than 90%. For the analysis of S6 kinase activity half of each sample was removed and adjusted to 50  $\mu$ M ATP and 10 mM MgCl<sub>2</sub>, and then incubated at 30°C for 15 min with either MAP kinase or buffer. Then 20  $\mu$ g of 40S ribosomal subunits and [ $\gamma$ - $^{32}$ P]ATP (final concentration 150  $\mu$ M; approximately  $4 \times 10^3$  c.p.m. pmol<sup>-1</sup>) were added and incubation was continued for 15 min at 30°C. S6 K II without MAP kinase retained more than 70% of its activity after incubation under these conditions for 15 min. Upon further incubation in the presence of ATP and Mg<sup>2+</sup> more than 80% of the residual activity was retained and, in the case of phosphatase-treated S6 K II, more than 90% of the residual activity was retained. For analysis of phosphorylation of S6 K II, the remainder of each sample was adjusted to 10 mM MgCl<sub>2</sub>, and incubated as described above in the presence of 123  $\mu$ M [ $\gamma$ - $^{32}$ P]ATP ( $58 \times 10^3$  c.p.m. pmol<sup>-1</sup>) without or with MAP kinase. For phosphopeptide mapping, the radiolabelled S6 K II was eluted from a preparative gel, precipitated, and digested with trypsin<sup>23</sup>. The phosphopeptides were fractionated on cellulose plastic-backed sheets by electrophoresis at pH 3.5 (pyridine:acetic acid:H<sub>2</sub>O, 1:10:189) in the first (horizontal) dimension and by ascending chromatography in the second dimension (bottom to top) in butanol:acetic acid:pyridine:H<sub>2</sub>O (15:3:10:12).

Figure 1 shows that partially purified MAP kinase was able to reactivate S6 K II that had been deactivated by preincubation with either protein phosphatase 1 or protein phosphatase 2A. Treatment of S6 K II with phosphatase 2A reduced S6 kinase activity by 98% (Fig. 1*a*, lanes 3 and 5), whereas phosphatase 1 reduced activity by 60% (Fig. 1*a*, lanes 7 and 9) as reported previously<sup>10</sup>. MAP kinase restored activity to 30–60% of the control (Fig. 1*a*, lanes 5, 6 and 9, 10). Even S6 K II that had not been incubated with phosphatase was activated nearly two-fold by MAP kinase (Fig. 1, lanes 3, 4 and 7, 8). Activation was dependent on the amount of MAP kinase, and no effect was observed with MAP kinase that had been heated to 80°C for 2 min. (data not shown). Significantly, the reactivation of S6 K II was associated with phosphorylation of the enzyme to the extent that the  $^{32}$ P content of S6 K II was significantly greater than that observed for autophosphorylated S6 K II (compare Fig. 1*b*,

lanes 3 and 4, 7 and 8). The stoichiometry of phosphorylation of S6 K II under maximal reactivation conditions was approximately 1:2 (phosphate: protein molecules) (data not shown). Similar results were obtained in multiple experiments in both laboratories with two different S6 K II preparations and two different MAP kinase preparations.

Since S6 K II undergoes autophosphorylation (ref. 1 and Fig. 1*b*, lanes 3 and 7) it was important to rule out the possibility that MAP kinase merely stimulated autophosphorylation of S6 K II. As shown in Fig. 1*c*, phosphorylation of S6 K II by MAP kinase resulted in the appearance of at least two new phosphopeptides regardless of whether S6 K II had been incubated with phosphatase 1. Similar results were obtained after deactivation by phosphatase 2A. This indicates that there is at least one phosphorylation site for MAP kinase that is distinct from autophosphorylation sites. The similarity in migration of



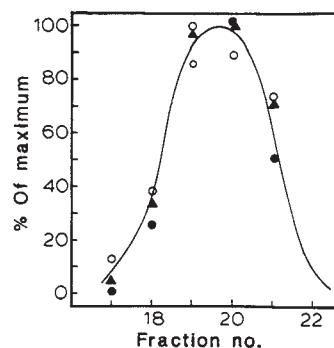


**Fig. 2** Insulin-dependence of S6 K II reactivation. S6 K II was treated with either protein phosphatase 2A or protein phosphatase 1 in the presence or absence of the appropriate inhibitor and then incubated in the presence of ATP and either buffer or MAP kinase prepared from resting or insulin-treated 3T3-L1 cells as described<sup>12</sup>. Then 40S subunits and [ $\gamma$ -<sup>32</sup>P]ATP were added for determination of S6 kinase activity as described in the legend to Fig. 1. All data shown are corrected for c.p.m. in S6 when the MAP kinase preparations or MAP kinase buffer were incubated alone with 40S subunits and [ $\gamma$ -<sup>32</sup>P]ATP. After this correction, the incorporation of radiolabel into S6 in samples in which MAP kinase buffer was used to reactivate was used as the 100% (control) value for the identical samples reactivated with the resting cell or insulin-stimulated MAP kinase preparations. In these experiments the average deactivation was 98% and 73% for protein phosphatase 2A and 1, respectively. Data are expressed as the mean  $\pm$  standard error of the mean ( $n = 5$  for phosphatase 2A and  $n = 3$  for phosphatase 1).

the new phosphopeptides for S6 K II phosphorylated by MAP kinase with and without phosphatase pretreatment suggests MAP kinase phosphorylates the same sites in both cases. Consistent with this idea, phosphoamino acid analysis of S6 K II phosphorylated by MAP kinase with or without phosphatase treatment resulted in the *de novo* appearance of phosphothreonine (Fig. 1d). Since S6 K II autophosphorylation occurs exclusively on serine residues (Fig. 1d, lane 4) at least one MAP kinase phosphorylation site in S6 K II is a threonine residue.

Previous studies have shown that insulin treatment of 3T3-L1 cells results in a 20-fold elevation of MAP kinase activity<sup>12</sup>. Only a preparation from insulin-treated cells, which contained elevated MAP kinase activity, was able to reactivate S6 K II that had been deactivated by either protein phosphatase 1 or protein phosphatase 2A (Fig. 2). This indicates that the activity responsible for reactivation is insulin stimulated and is specifically associated with elevated MAP kinase activity. The extensive reactivation by MAP kinase of S6 K II after deactivation by protein phosphatase 2A, relative to phosphatase 1 (Fig. 2), reflects the greater degree of inactivation of S6 K II by phosphatase 2A (98% Fig. 1a, lanes 3 and 5) relative to that by phosphatase 1, (60–80%, Fig. 1a, lanes 7 and 9, and Fig. 2). Other experiments showed that reactivated S6 K II could be deactivated again by a second phosphatase treatment (data not shown).

To obtain further evidence that the MAP kinase was responsible for S6 K II reactivation, another preparation of partially purified MAP kinase was subjected to an additional purification step by gel filtration on Superose 12 (ref. 13), and individual fractions were assayed for MAP kinase activity, phosphorylation and reactivation of dephosphorylated S6 K II. At this stage of purification, the 42 kDa MAP kinase was the major protein in



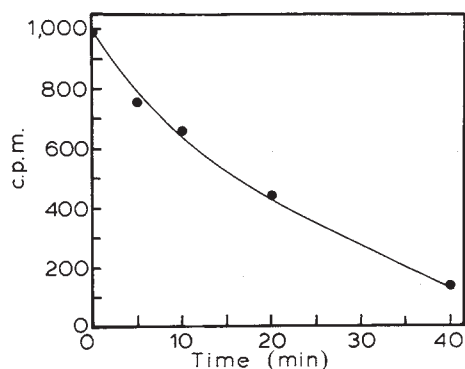
**Fig. 3** Co-elution of MAP kinase (●) and activity phosphorylating (○) and reactivating (▲) S6 K II upon Superose 12 gel filtration chromatography. MAP kinase was applied to a Superose 12 FPLC column (Pharmacia) equilibrated with buffer A (50 mM HEPES, pH 7.5 at 5 °C, 50 mM NaCl, 10% glycerol, 1 mM DTT). Fractions were assayed for kinase activity using either MAP-2 (●) or S6 K II (○) as substrate. Activity of S6 K II (▲) was assayed using 40S ribosomal subunits as substrate. The data are expressed as per cent of maximal incorporation of <sup>32</sup>P which was measured by liquid scintillation spectroscopy of MAP-2 (2670 c.p.m.) or ribosomal protein S6 (1002 c.p.m.) excised from polyacrylamide gels, or by densitometry of autoradiograms of dried silver-stained gels containing S6 K II.

the preparation<sup>13</sup>. As shown in Fig. 3, all three activities co-eluted from the gel filtration column. This provides strong evidence that the activation of S6 K II is mediated by MAP kinase.

Previous studies have shown clearly that MAP kinase is activated by insulin in 3T3-L1 cells just prior to S6 kinase activation<sup>12</sup>. Although the results shown here indicate that S6 kinase is activated via phosphorylation by MAP kinase, little is known about how MAP kinase itself is activated. However, it is of interest that, when isolated from <sup>32</sup>P-labelled, insulin-treated cells, purified MAP kinase contains both phosphothreonine and phosphotyrosine<sup>14</sup>. MAP kinase is thus the first insulin target enzyme demonstrated to contain phosphotyrosine *in vivo* after insulin stimulation. To determine whether these phosphorylations might play a role in MAP kinase activity, the enzyme was incubated with one of the protein phosphatases and assayed for MAP kinase activity. Incubation with protein phosphatase 2A resulted in a time-dependent inactivation of MAP kinase activity (Fig. 4), whereas phosphatase 1 had no effect (data not shown). Although phosphatase 1 is absolutely specific for dephosphorylation of serine and threonine residues, phosphatase 2A has significant activity against phosphotyrosyl sites as well<sup>15</sup>. Additional experiments will be required to determine whether the selective ability of phosphatase 2A to deactivate MAP kinase is due to its selective ability to dephosphorylate phosphotyrosyl sites. However, as yet, we have been unable to reactivate MAP kinase with any of several tyrosine kinases, including the insulin receptor kinase.

Our results demonstrate the reconstitution *in vitro* of one step of a protein kinase cascade initiated *in vivo* by insulin in an unidentified manner. Although the purified MAP kinase preparation is not homogenous, several lines of evidence argue strongly that this activity is responsible for S6 kinase activation: MAP kinase is activated *in vivo* by insulin just prior to S6 kinase activation<sup>12</sup>; both MAP kinase and the factor activating S6 K II are dramatically stimulated in partially-purified fractions from cells treated with insulin (ref. 13 and Fig. 2); both activities co-purify during ion exchange, hydrophobic interaction and gel filtration chromatography (Fig. 3); and unique phosphopeptides and phosphoamino acids appear in S6 K II after phosphorylation by MAP kinase (Fig. 1c and d).

Although the concept of protein kinase cascades dominates the field of protein phosphorylation, the only previous examples



**Fig. 4** Inactivation of MAP kinase by dephosphorylation with protein phosphatase 2A. Portions of the pooled peak fractions of MAP kinase eluted from a Superose 12 column (legend, Fig. 3) were incubated with protein phosphatase 2A (10 U ml<sup>-1</sup> final) at 30 °C in the presence of 1 µg ml<sup>-1</sup> leupeptin, 2 µM pepstatin, 200 U ml<sup>-1</sup> aprotinin, and 0.1 mM phenylmethylsulphonyl fluoride. At the indicated times, the reaction was stopped by the addition of 10 mM KF and 1 mM EDTA. Aliquots were then assayed for MAP kinase activity. Results are expressed relative to controls incubated for the same period of time with phosphatase pretreated with KF and EDTA. The MAP kinase preparation itself is quite stable at this stage of purification and loses less than 30% of its activity during incubation at 30 °C for 40 min with inactivated phosphatase.

of a protein-serine kinase phosphorylating and activating another protein-serine kinase are cAMP-dependent protein kinase<sup>16</sup> and hydroxy-methyl-glutaryl CoA-reductase kinase<sup>17</sup>. The data reported here support the hypothesis that signal transduction by insulin also involves sequential activation by phosphorylation of a series of protein kinases. The approach we used to obtain these results was first to define the control of a biochemical process involved in the final cellular response: in this case to show that activation of an S6 kinase is responsible for the phosphorylation of S6. Then, after purification of the S6 kinase, it will be possible to work backwards to reveal how S6 kinase is activated by the insulin receptor. The present data indicate that the insulin signalling pathway involves the indirect activation of S6 kinase by phosphorylation by another soluble serine/threonine kinase, MAP kinase. Another step in the signal transduction pathway will be revealed when the method of activation of MAP kinase, which almost certainly involves its phosphorylation, becomes clear. Since S6 phosphorylation is stimulated in *Xenopus* oocytes by external insulin as well as by injected insulin receptor-kinase<sup>18</sup>, pp60<sup>v-src</sup> (ref. 19) and Abelson tyrosine kinase<sup>20</sup>, it may eventually be possible to determine whether MAP kinase is directly stimulated by different tyrosine kinases in injected oocytes or whether different mitogenic signals converge at an even earlier point in the S6 kinase activation pathway. Such investigations should yield a greater understanding of the importance of tyrosine protein kinase activity in the mechanism of action of insulin and many transforming proteins.

We thank Dr Jay Johansen for providing inhibitor-2 and purified catalytic subunits of protein phosphatases 1 and 2A, Gail Garner for technical assistance and Karen Eckart for secretarial support. This work was supported by grants from the NIH (J.L.M.), the American Cancer Society (J.L.M. and T.W.S.), and the Juvenile Diabetes Foundation (L.B.R.).

Received 5 January; accepted 28 June 1988.

1. Erikson, E. & Maller, J. L. *J. biol. Chem.* **261**, 350-355 (1986).
2. Erikson, E., Stefanovic, D., Blenis, J., Erikson, R. L. & Maller, J. L. *Molec. cell Biol.* **7**, 3147-3155 (1987).
3. Jones, S. W., Erikson, E., Blenis, J., Maller, J. L. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3377-3381 (1988).
4. Novak-Hofer, I. & Thomas, G. *J. biol. Chem.* **259**, 5995-6000 (1984).
5. Blenis, J. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7621-7625 (1985).

6. Tabarini, D., Heinrich, J. & Rosen, O. M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4369-4373 (1985).
7. Stefanovic, D., Erikson, E., Pike, L. J. & Maller, J. L. *EMBO J.* **5**, 157-160 (1986).
8. Cobb, M. H. *J. biol. Chem.* **261**, 12994-12999 (1986).
9. Nemenoff, R. A., Gunsalus, J. & Avruch, J. *Archs Biochem. Biophys.* **245**, 196-203 (1986).
10. Maller, J. L. *J. Cyclic Nuc. Prot. Phos. Res.* **11**, 543-555 (1987).
11. Ballou, L. M., Jen, P. & Thomas, G. *J. biol. Chem.* **263**, 1188-1194 (1988).
12. Ray, L. B. & Sturgill, T. W. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1502-1506 (1987).
13. Ray, L. B. & Sturgill, T. W. *J. biol. Chem.* **263**, (in the press).
14. Sturgill, T. W. & Ray, L. B. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3753-3757 (1988).
15. Foulkes, J. G. *Curr. Top. Microbiol. Immunol.* **107**, 163-180 (1983).
16. Walsh, D. A., Perkins, J. P. & Krebs, E. G. *J. biol. Chem.* **243**, 3763-3765 (1968).
17. Gibson, D. M., Parker, R. A., Stewart, C. S. & Evenson, K. J. *Adv. Enzyme Regul.* **20**, 263-283 (1982).
18. Maller, J. L. *et al. Nature* **320**, 459-461 (1986).
19. Spivack, J. G., Erikson, R. L. & Maller, J. L. *Molec. cell Biol.* **4**, 1631-1634 (1984).
20. Maller, J. L., Foulkes, J. G., Erikson, E. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **82**, 272-276 (1985).
21. Tung, H. Y. L., Resink, T. J., Hemmings, B. A., Shenolikar, S. & Cohen, P. *Eur. J. Biochem.* **138**, 635-641 (1984).
22. Tonks, N. K. & Cohen, P. *Eur. J. Biochem.* **145**, 65-70 (1984).
23. Erikson, E., Brugge, J. S. & Erikson, R. L. *Virology* **80**, 177-185 (1977).

## Abolition of actin-bundling by phosphorylation of human erythrocyte protein 4.9

Athar Husain-Chishti, Andrew Levin & Daniel Branton

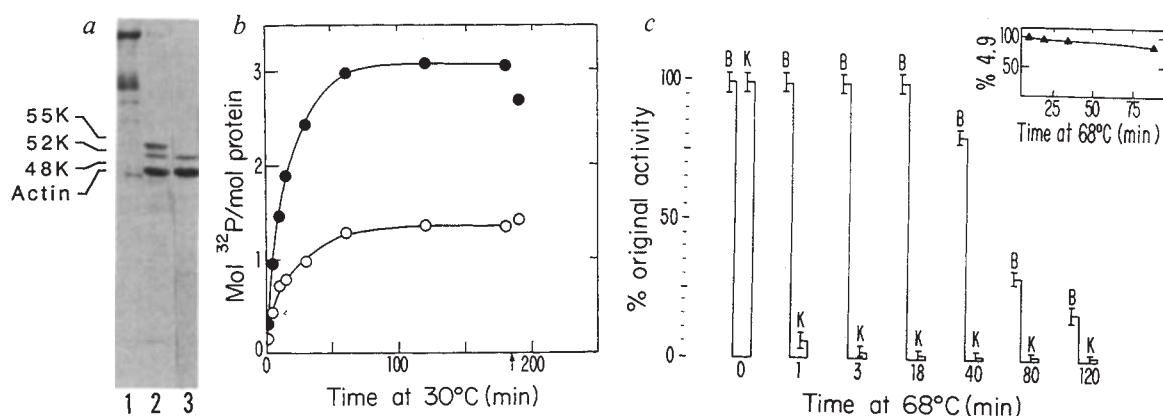
Department of Cellular and Developmental Biology, Biological Laboratories, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

**Protein 4.9, first identified as a component of the human erythrocyte membrane skeleton, binds to and bundles actin filaments<sup>1</sup>. Protein 4.9 is a substrate for various kinases, including a cyclic AMP(cAMP)-dependent one, *in vivo* and *in vitro*<sup>2-7</sup>. We show here that phosphorylation of protein 4.9 by the catalytic subunit of cAMP-dependent protein kinase reversibly abolishes its actin-bundling activity, but phosphorylation by protein kinase C has no such effect. A quantitative immunoassay showed that human erythrocytes contain 43,000 trimers of protein 4.9 per cell, which is equivalent to one trimer for each actin oligomer in these red blood cells. As analogues of protein 4.9 have been identified together with analogues of other erythroid skeletal proteins in non-erythroid tissues of numerous vertebrates<sup>8</sup>, phosphorylation and dephosphorylation of protein 4.9 may be the basis for a mechanism that regulates actin bundling in many cells.**

Purified human erythrocyte protein 4.9 is a trimer composed of two polypeptides of relative molecular mass ( $M_r$ ) 48,000 (48K) and 52,000 (52K)<sup>1</sup> (Fig. 1a). The amino acid compositions of these two polypeptides show some differences, but one- and two-dimensional peptide mapping and antibody cross-reactivity show that the two polypeptides have extensive homology (A.H.-C. *et al.*, manuscript in preparation). Both polypeptides bind to, and sediment with, actin bundles<sup>1</sup> (Fig. 2, inset). Cross-linking experiments have shown that both isoforms participate in trimer formation (A.H.-C. *et al.*, manuscript in preparation). In solutions of the purified protein, the two isoforms were present in a ratio of approximately three 48K polypeptides to one 52K polypeptide and this ratio was essentially identical in intact cells and FPLC-purified protein 4.9 (Fig. 1a, lane 3).

For convenience, we refer to protein 4.9 isolated from normal cells as the 'dephospho-form', even though direct phosphate analysis indicated that it contained 0.45 mole of phosphate per subunit. When maximally phosphorylated by the catalytic subunit of cAMP-dependent protein kinase, the 48K polypeptide incorporated an additional 3 mole of phosphate per mole of subunit, and the 52K polypeptide incorporated an additional 1.5 mole of phosphate per mole of subunit (Fig. 1b). No proteolysis of protein 4.9 was detected during incubation with the protein kinase (not shown). The actin-bundling activity of protein 4.9 was stable at 68 °C for at least 18 min, whereas 3 min at this temperature completely inactivated the catalytic subunit



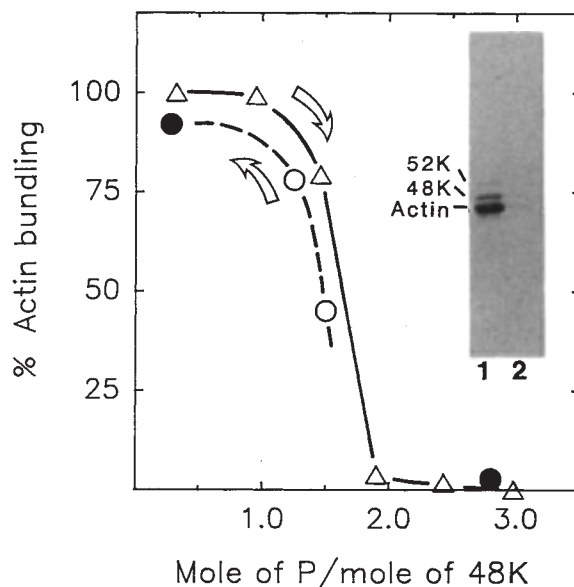


**Fig. 1** Purification, phosphorylation and heat stability of protein 4.9. *a*, Purification of protein 4.9. Lane 1, sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) of erythrocyte ghosts; lane 2, partially purified protein 4.9 after chromatography on DEAE-Sephacel; lane 3, purified protein 4.9 after chromatography on a Mono Q column. *b*, Phosphorylation of protein 4.9 by the catalytic subunit of cAMP-dependent protein kinase. Closed circles, 48K polypeptide; open circles, 52K polypeptide. Arrow indicates the addition of 3 mM ATP of the same specific activity as initially to test for substrate depletion. *c*, The effect of temperature on the actin-bundling activity of protein 4.9 (B), the kinase activity of the cAMP-dependent protein kinase (K), and the amount of protein 4.9 that remained in solution during heating (inset).

**Methods.** Protein 4.9 (300  $\mu\text{g ml}^{-1}$ ) or kinase subunit (57  $\mu\text{g ml}^{-1}$ ) in a buffer containing 20 mM Tris-HCl, pH 7.5, 1 mM EGTA, 0.5 mM DTT, 0.02% sodium azide and 120 mM NaCl was incubated at 68 °C. Samples were taken at different times and assayed either for actin-bundling activity or kinase activity. Actin-bundling was assayed in a mixture (50  $\mu\text{l}$ ) containing 55  $\mu\text{g ml}^{-1}$  protein 4.9, 280  $\mu\text{g ml}^{-1}$  actin 50 mM Tris-HCl, pH 7.5, 2 mM  $\text{MgCl}_2$ , 1 mM ATP, 0.04 mM  $\text{CaCl}_2$ , 0.7 mM EGTA, 1 mM DTT, 75 mM NaCl and 0.02% azide; the contents were incubated for 2 h at 30 °C and subsequently layered over 150  $\mu\text{l}$  of 15% cold sucrose. After centrifugation for 15 min at 12,000g in a swinging bucket rotor, the supernatant was carefully removed, sample buffer was added to the pellet, heated and subjected to SDS-PAGE<sup>22</sup>. The Coomassie-stained bands were excised and quantified<sup>23</sup>. Protein kinase activity was assayed using protein 4.9 as substrate<sup>7</sup>. Protein 4.9 precipitation at 68 °C was assayed by measuring protein loss in the supernatant of samples taken after centrifugation at 150,000g for 30 min in an airfuge.

**Fig. 2** The effect of phosphorylation on the actin-bundling activity of protein 4.9. Phosphorylation (triangles) abolishes actin-bundling activity; dephosphorylation (circles) restores activity. Inset: the effect of phosphorylation on bundles previously generated by dephosphorylated protein 4.9. Lane 1, proteins sedimented when actin bundles were treated with 30  $\mu\text{g ml}^{-1}$  heat-inactivated kinase catalytic subunit; lane 2, as lane 1 but treated with active (not heat treated) kinase catalytic subunit for 15 min. The amount of actin seen in lane 2 is not significantly greater than the amount seen in equivalent samples containing no protein 4.9.

**Methods.** Phosphorylation samples (40  $\mu\text{l}$  each, triangles) were taken at timed intervals from a reaction stock parallel to that described in Fig. 1b and incubated at 68 °C for 3 min to inactivate protein kinase. Actin bundling was initiated by the addition of monomeric actin. The assay mixture (50  $\mu\text{l}$ ) contained 240  $\mu\text{g ml}^{-1}$  actin, 85  $\mu\text{g ml}^{-1}$  protein 4.9, 12  $\mu\text{g ml}^{-1}$  kinase catalytic subunit. Actin bundling by protein 4.9 and phosphorylation of the 48K subunit were quantified as described in Fig. 1. The phosphorylation values represent moles of labelled phosphate incorporated into the dephospho-form. Dephosphorylation samples were taken from independent experiments using either 90  $\mu\text{g ml}^{-1}$  Sigma type XXX alkaline phosphatase (open circles) or 700 units  $\text{ml}^{-1}$  type 1 protein phosphatase<sup>24</sup> (solid circles). In either case, the enzyme, in a 40  $\mu\text{l}$  assay volume containing 20 mM Tris-HCl, pH 8.0 and 3  $\mu\text{g ml}^{-1}$  each of leupeptin, pepstatin, antipain and chymostatin (and 2 mM  $\text{MnCl}_2$  in the case of type 1 phosphatase), was incubated at 30 °C with 72  $\mu\text{g ml}^{-1}$   $^{32}\text{P}$ -labelled protein 4.9 that had been maximally phosphorylated by kinase catalytic subunit. After incubation, actin-bundling by protein 4.9 was quantified as in Fig. 1; dephosphorylation was quantified by densitometry of autoradiographs relative to maximally phosphorylated protein 4.9 (ref. 7).

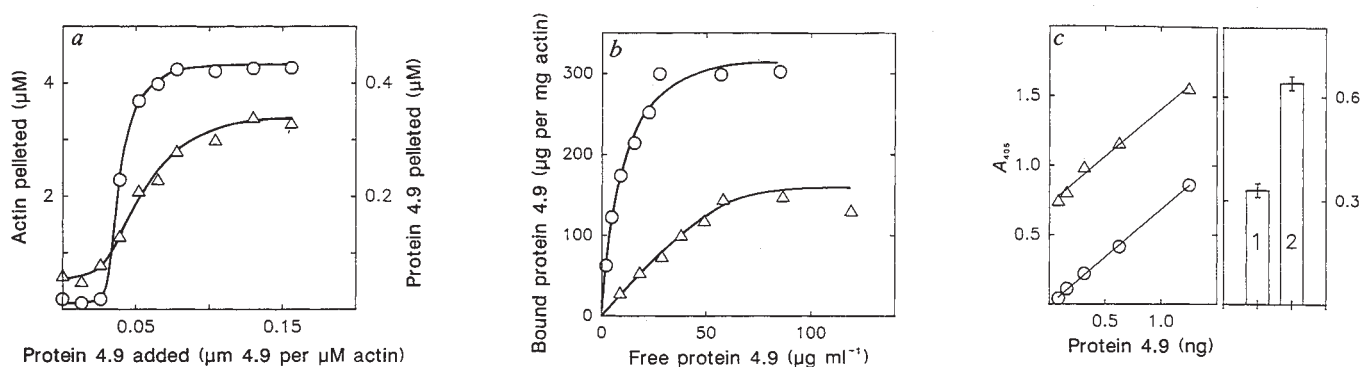


of cAMP-dependent protein kinase (Fig. 1c). This differential temperature sensitivity made it possible to measure the effect of phosphorylation on the actin-bundling activity without removing the kinase catalytic subunit from the assay mixture (Fig. 2).

The actin-bundling activity of protein 4.9 was inhibited by phosphorylation (Fig. 2, triangles). Two moles of phosphate per mole of the 48K subunit inhibited all bundling activity. This inhibition was reversible: when fully phosphorylated protein 4.9 was dephosphorylated with phosphatase, bundling activity was restored (Fig. 2, circles). Actin which had been bundled by

dephosphorylated protein 4.9 was unbundled in the presence of exogenous catalytic subunit under conditions favourable for protein 4.9 phosphorylation (Fig. 2, inset). The abolition of actin-bundling activity was not due to limiting amounts of protein 4.9: the effects of phosphorylation were observed at 0.06 to 0.15 mole of trimeric protein 4.9 incubated per mole of actin, a range that includes more than the amount of protein 4.9 needed for optimal bundling (Fig. 3a).

Phosphorylation by protein kinase C incorporated 0.61 and 0.24 mole of phosphate per mole of the 48K and 52K subunits of protein 4.9, respectively (data not shown). This did not



**Fig. 3** Actin bundling and binding. *a*, Quantification of actin bundling by dephospho-protein 4.9. Actin in pellet, circles; protein 4.9 in pellet, triangles. A constant amount (4.82 μM) of actin was bundled by increasing amounts of protein 4.9 in buffer conditions described in Fig. 1. Supernatants and pellets were carefully separated after 12,000g centrifugation for 20 min (no actin filaments sedimented at this speed), analysed by SDS-PAGE and quantified<sup>23</sup>. A  $M_r$  of 150K was assumed for protein 4.9. *b*, Effect of phosphorylation on the binding of protein 4.9 to actin filaments. Dephospho-protein 4.9, circles; maximally phosphorylated protein 4.9, triangles. Phosphorylation and assay conditions were as in Fig. 1. Actin filaments were pelleted at 155,000g for two hours in a type 42.2 Ti rotor and pellets were solubilized in 1% SDS sample buffer. The protein 4.9 content of these solubilized pellets was determined using a quantitative immunoassay (see below). Data points represent an average of duplicate determinations which were within 5%. *c*, Immunoassay standardization and validation. Left: to test for any effect of solubilized crude erythrocyte ghosts on the standard curve using purified protein 4.9 (circles), a constant amount of solubilized plasma membranes (containing sufficient protein 4.9 to produce 0.7 absorbance units in the immunoassay) was added to individual wells in a 96-well microtitre plate together with increasing amounts of FPLC-purified protein 4.9 (triangles). Duplicate points were within 2%. Right: the amount of protein 4.9 in 5 μl (1) or 10 μl (2) of a sample of ghosts (about 0.45 μg protein ml<sup>-1</sup>). Doubling the plasma membrane concentration precisely doubled the absorbance values.

**Methods.** FPLC-purified protein 4.9 was injected into BALB/c female mice and fusion with FoxNY myeloma cells was carried out by standard procedures<sup>25</sup>. Hybridomas were selected by ELISA and were subcloned by serial dilution. Two monoclonal antibodies (both of IgG<sub>1</sub> subisotype by double diffusion) specific for non-overlapping epitopes of protein 4.9 (data not shown) were used. One monoclonal antibody in phosphate-buffered saline (250 ng per well) was bound to plastic 96-well Immulon 2 microtitre plates with flat bottoms (Dynatech Laboratories) and served to capture soluble protein 4.9; captured protein 4.9 was then detected by incubation with the second, biotinylated, monoclonal antibody (100 ng per well) followed by streptavidin-alkaline phosphatase. The hydrolysis of the phosphatase substrate p-nitrophenyl phosphate was quantified in microtitre wells at 405 nm. Both monoclonal antibodies were characterized with respect to dephosphorylated and phosphorylated forms of protein 4.9 and were found to be similar in their reactivities against the two forms. Standard protein 4.9 concentration was determined by a micro Kjeldahl procedure<sup>26</sup>. To quantify protein 4.9, samples were solubilized in 1% SDS and diluted 200-fold with blocking buffer (phosphate-buffered saline, 5% neonatal calf serum, 0.1% Tween-20 and 0.02% azide). FPLC-purified protein 4.9 samples were supplemented with 0.005% SDS and used as a standard. The number of erythrocyte ghosts was determined by Coulter Counter ZM and verified by microscopy in a haemocytometer counting chamber.

diminish the actin-bundling activity of protein 4.9. We do not know whether this is due to the relatively small amount of phosphate incorporated into protein 4.9 by protein kinase C or to the fact that protein kinase C phosphorylates different sites on protein 4.9 than does the cAMP-dependent kinase (ref. 3 and A.H.-C. *et al.*, manuscript in preparation).

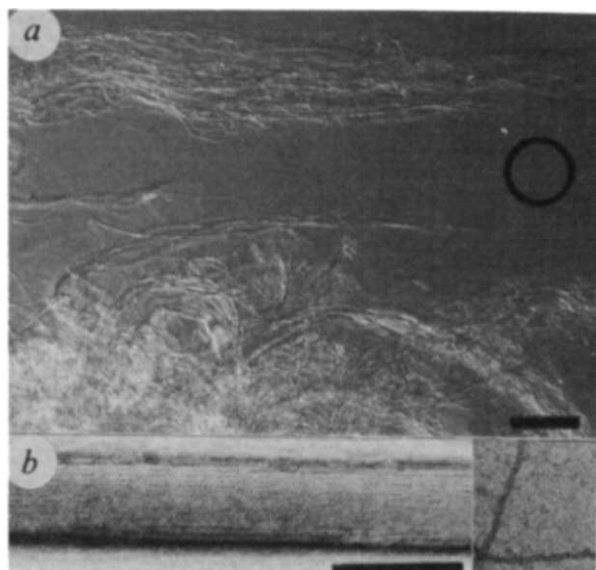
Quantitative evaluation of actin-bundling by dephospho-protein 4.9 indicated that detectable bundling was observed with as little as 0.04 mole of protein 4.9 per mole of actin (Fig. 3a). At saturation, protein 4.9 and actin in bundles showed a stoichiometry of one trimer per  $12 \pm 1.5$  actin monomers (average  $\pm$  standard deviation,  $N=3$ ). High-speed centrifugation that sedimented both bundles and actin filaments showed that both the active dephospho-form of protein 4.9 as well as the maximally phosphorylated inactive form bound and sedimented with F-actin (Fig. 3b). The maximal binding stoichiometry of protein 4.9 to actin differed for the two forms: one trimer per  $11.8 \pm 0.4$  actin monomers for the dephospho-form and one trimer per  $23 \pm 5$  actin monomers (average  $\pm$  standard deviation,  $N=3$ ) for the fully phosphorylated form. Scatchard analysis using a single-binding-site model and a  $M_r$  of 150K for protein 4.9 gave  $K_d$  values of 57 nM and 280 nM for the dephosphorylated and phosphorylated forms of protein 4.9, respectively, but the validity of such an analysis is questionable because of the changing geometry and steric hindrance generated as actin is cross-linked to form actin bundles.

Because both the dephospho- and the phosphorylated forms of protein 4.9 migrated as a trimeric particle (Stokes radius of 4.8 nm on a Sephacryl S-100 column), each of the subunits of protein 4.9 need have only one binding site for the whole trimeric protein to be an active bundling factor. In fact, analysis of the association between protein 4.9 and actin suggests a more com-

plicated situation, akin to that seen in actin-bundling proteins that are monomers in solution. Such proteins appear to bundle actin because they have two (or more) actin-binding sites per polypeptide<sup>9</sup>. If each subunit has two binding sites, the association of protein 4.9 subunits into trimers should not be required for bundling activity; the bivalency (or multivalency) in single subunits may be sufficient to induce bundling. To test this possibility, we incubated F-actin with protein 4.9 in mixtures containing 50–70% glycerol, concentrations that have been observed to promote the dissociation of protein 4.9 subunits<sup>1</sup>. We verified the dissociation of trimeric protein 4.9 into its subunits by gel exclusion chromatography (Superose 12 column using Pharmacia FPLC): at 50% and higher concentrations of glycerol, protein 4.9 migrated as a monomer with a Stokes radius of 3.1 nm. Although bundling was slow in 50 or 70% glycerol (one to two days at 4 °C) and quantification was not possible because actin could not be sedimented at these high glycerol concentrations, dephospho-protein 4.9 produced bundles that were visible in the light microscope and appeared similar to those generated by the native trimers (Fig. 4a). Although we cannot rule out the possibility that F-actin promotes or stabilizes protein 4.9 trimers in high glycerol concentrations, the simplest interpretation is that two or more binding sites on monomeric protein 4.9 are sufficient to bundle actin when in the dephospho-form. If this interpretation is correct, protein 4.9 fails to bundle actin when phosphorylated because some binding sites are lost or because the binding sites no longer retain the actin filaments in the correct parallel orientation required for bundling.

An immunoassay was used to quantify the total amount of protein 4.9 per ghost. The immunoassay provided a good measure of pure protein 4.9 at concentrations from 0.2 ng ml<sup>-1</sup> to 80 ng ml<sup>-1</sup> (not shown) and was unaffected by solubilized





**Fig. 4** *a*, Differential interference contrast micrograph of rabbit muscle actin incubated with dephospho-protein 4.9 in the buffer conditions of Fig. 1. Similar bundles were observed when the buffer was supplemented with 50% or 70% glycerol. The inset circle is a representation of the empty fields seen when actin filaments were incubated with phosphorylated protein 4.9. *b*, Electron micrograph of rabbit muscle actin bundled by dephospho-protein 4.9 with inset, at same magnification, showing control actin filaments. Scale bars, 10  $\mu\text{m}$  (*a*) and 0.1  $\mu\text{m}$  (*b*).

crude plasma membranes (Fig. 3c). Using three different ghost preparations containing between 0.02 and 0.2  $\mu\text{g ml}^{-1}$  membrane protein, we measured  $43,300 \pm 900$  (average  $\pm$  standard deviation,  $N=3$ ) trimers of protein 4.9 per ghost. No protein 4.9 was detectably lost to the cytosolic fractions during ghost production. As actin in human erythrocytes is assembled into extraordinarily regular oligomers, each of which appears to contain 12–13 monomers<sup>10</sup>, and since there are 500,000 monomers per human erythrocyte<sup>11</sup>, there must be 40,000 of these oligomeric junctions per cell. Our results show that there are 43,000 trimers of protein 4.9 per erythrocyte, and that in actin bundled by saturating quantities of protein 4.9 there is one trimer associated with 11–13 actin monomers. These results indicate that one protein 4.9 is bound to each actin oligomer in the mature erythrocyte. The maximum distance between oligomers is limited by the 200 nm contour length of the spectrin tetramers that cross-link the oligomers into the membrane skeleton<sup>12</sup>. Because the spectrin chains are flexible, the distance between oligomers could in principle be anywhere from 0 to 200 nm in the intact cell. Thus, if *in vivo* dephospho-protein 4.9 bundles the actin oligomers, it could have a profound effect on the physical properties of the membrane skeleton and the shape of the erythrocyte<sup>12</sup>. To elucidate further the role of protein 4.9, it will be important to determine how the stoichiometry of protein 4.9 to actin changes during erythropoiesis and whether the distribution of actin and protein 4.9 is altered by phosphorylation-dephosphorylation reactions *in vivo*.

Although the large variety of actin-binding proteins expressed in different cells plays a critical role in modulating actin function, relatively little is known of how the interactions of these proteins with actin are regulated<sup>9,13</sup>. Our work shows that erythrocyte protein 4.9 is a major substrate for cAMP-dependent protein kinase and loses its actin-bundling activity when phosphorylated. These properties are shared by synapsin I, the neuron-specific phosphoprotein<sup>14</sup>, but these two proteins are not otherwise thought to be related. In fact, synapsin I has been proposed as a brain analogue of erythrocyte protein 4.1<sup>15</sup>, a protein that

is found in many different cell types<sup>16,17</sup> and that is known to stabilize the interaction of spectrin with actin<sup>18,19</sup>. In nervous tissue, protein 4.9 is localized primarily to the cell body of neurons<sup>8</sup>, whereas synapsin I is localized primarily to the cytoplasmic surface of synaptic vesicles<sup>20,21</sup>. Thus, analogues of protein 4.9 may coexist with analogues of synapsin I and participate in phosphorylation-dephosphorylation events which modulate actin filament arrays and the organization of the cytoskeleton.

We thank Mr Chi-Chih Wu for assistance with the peptide mapping and gel filtration, Drs Thomas Pesacreta and Cam Owen for microscopy and Mr William Faquin for discussions. This work was supported by the NIH and an American Heart Association (Massachusetts Branch) postdoctoral fellowship to A.H.-C.

Received 15 May 1987; accepted 5 July 1988.

1. Siegel, D. L. & Branton, D. *J. Cell Biol.* **100**, 775–785 (1985).
2. Ling, E. & Sapirstein, V. *Biochem. biophys. Res. Commun.* **120**, 291–298 (1984).
3. Horne, W. C., Leto, T. L. & Marchesi, V. T. *J. Biol. Chem.* **260**, 9073–9076 (1985).
4. Palfrey, H. C. & Waseem, A. *J. Biol. Chem.* **260**, 16021–16029 (1985).
5. Faquin, W. C., Chahwala, S. B., Cantley, L. C. & Branton, D. *Biochim. biophys. Acta* **887**, 142–149 (1986).
6. Cohen, C. M. & Foley, S. F. *J. Biol. Chem.* **261**, 7701–7709 (1986).
7. Husain, A. & Branton, D. *Analyt. Biochem.* **155**, 206–211 (1986).
8. Faquin, W. C., Husain, A., Hung, J. & Branton, D. *Eur. J. Cell Biol.* **46**, 168–175 (1988).
9. Pollard, T. D. & Cooper, J. A. *Rev. Biochem.* **55**, 987–1035 (1986).
10. Byers, T. J. & Branton, D. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6153–6157 (1985).
11. Pinder, J. C. & Gratzner, W. B. *J. Cell Biol.* **96**, 768–775 (1983).
12. Elgsaeter, A., Stokke, B. T., Mikkelsen, A. & Branton, D. *Science* **234**, 1217–1223 (1986).
13. Craig, S. W. & Pollard, T. D. *Trends biochem. Sci.* **7**, 88–92 (1982).
14. Bahler, M. & Greenhard, P. *Nature* **326**, 704–707 (1987).
15. Baines, A. J. & Bennett, V. *Nature* **315**, 410–413 (1985).
16. Cohen, C. M., Foley, S. F. & Korgsgren, C. *Nature* **299**, 648–650 (1982).
17. Spiegel, J. E., Beardsley, D. S., Southwick, F. S. & Lux, S. E. *J. Cell Biol.* **99**, 886–893 (1984).
18. Ungewickell, E., Bennett, P., Calvert, R., Ohanian, V. & Gratzner, W. *Nature* **280**, 811–814 (1979).
19. Fowler, V. & Taylor, D. L. *J. Cell Biol.* **85**, 361–376 (1980).
20. DeCamilli, P., Harris, J. M. Jr., Hunter, W. B. & Greengard, P. *J. Cell Biol.* **96**, 1355–1373 (1983).
21. Huttner, W. B., Schiebler, W. & Greengard, P. *J. Cell Biol.* **96**, 1374–1388 (1983).
22. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
23. Malloy, J. M., Rieker, J. P. & Rizzo, C. F. *Analyt. Biochem.* **141**, 503–509 (1984).
24. Ingebritsen, T. S. & Cohen, P. *Science* **221**, 331–338 (1983).
25. Cuello, A. C., Milstein, C. & Galfre, G. in *Immunohistochemistry* (ed. Cuello, A. C.) 215–256 (John Wiley & Sons, New York 1983).
26. Jaenicke, L. *Analyt. Biochem.* **61**, 623–627 (1974).

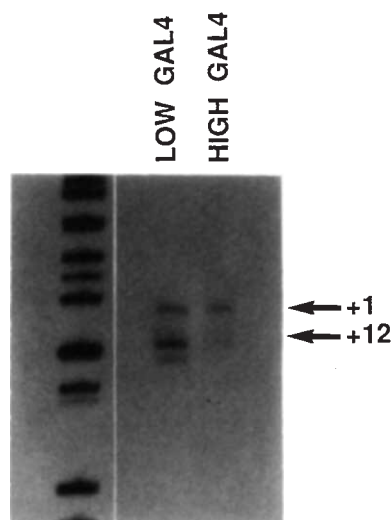
## Negative effect of the transcriptional activator GAL4

Grace Gill & Mark Ptashne

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138, USA

The yeast transcriptional activator GAL4 binds specific sites on DNA to activate transcription of adjacent genes<sup>1–5</sup>. The distinct activating regions of GAL4 are rich in acidic residues and it has been suggested that these regions interact with another protein component of the transcriptional machinery (such as the TATA-binding protein or RNA polymerase II) while the DNA-binding region serves to position the activating region near the gene<sup>6,7,8</sup>. Here we show that various GAL4 derivatives, when expressed at high levels in yeast, inhibit transcription of certain genes lacking GAL4 binding sites, that more efficient activators inhibit more strongly and that inhibition does not depend on the DNA-binding domain. We suggest that this inhibition, which we call squelching, reflects titration of a transcription factor by the activating region of GAL4.

GAL4 binds to four related sites in the GAL upstream activation sequence (UAS<sub>G</sub>) and stimulates transcription of the adjacent genes<sup>1–5</sup>. The DNA-binding and transcriptional activation functions of GAL4 are separable<sup>3,6</sup>. Deletion analysis of GAL4 defines two activating regions (region I, residues 148–236, and



**Fig. 1** High levels of GAL4 inhibit inducible HIS3 transcription. Primer extension analysis was performed on 30 µg total yeast RNA using a primer corresponding to positions +172 to +153 of the HIS3 gene<sup>29</sup>. RNA was isolated from the yeast strain DBY745 (adel-100, leu2-2,112, ura3-52) bearing a multicopy plasmid expressing GAL4 from the ADH1 promoter (pLKC15, ref. 30) or a control plasmid lacking GAL4 (Yep213). Cells were grown in minimal media lacking leucine and containing 2% galactose, 3% glycerol and 2% lactic acid, pH6. RNA was isolated and primer extensions performed as described<sup>9</sup>, except actinomycin D was included in the extension reaction at a final concentration of 50 µg ml<sup>-1</sup>. The transcripts are numbered according to Struhl<sup>20,29</sup>. The markers shown are pBR322 digested with *Msp*1.

region II, residues 768–881), either of which activates transcription when fused to the DNA-binding domain (residues 1–147)<sup>9</sup>. These activating regions, like that described for another yeast activator protein, GCN4, are rich in acidic residues, although they share no sequence homology<sup>9,10</sup>. Mutagenesis of activating region I of GAL4 reveals a strong correlation between the efficiency of transcriptional activation and the acidity of the region<sup>11</sup>. A functional activator is generated by replacing the activating regions of GAL4 with any of a large number of acidic sequences encoded by *Escherichia coli* genomic fragments or by a peptide designed to form an amphipathic  $\alpha$ -helix with acidic residues along one surface<sup>12,13</sup>. Here we show that at high concentrations GAL4 inhibits transcription of certain genes lacking the GAL4 binding site. We map this inhibitory effect to the activating region of GAL4 and show that derivatives bearing stronger activating regions inhibit more efficiently. We show that, of two closely related transcripts, only the one which is inducible by GAL4 when the UAS<sub>G</sub> is present<sup>14</sup> is inhibited by high levels of GAL4. Simultaneous expression at high levels of the negative regulator GAL80, a protein that masks the activating regions of GAL4 and inhibits its stimulatory activity<sup>15,16</sup>, relieves this inhibition.

High levels of GAL4 in yeast inhibit expression of a GAL1:*lacZ* fusion gene lacking GAL4 binding sites. Table 1 shows that GAL4, expressed from the strong alcohol dehydrogenase (ADH) promoter on a multicopy plasmid, greatly stimulates expression of a GAL1:*lacZ* fusion gene bearing UAS<sub>G</sub>. Expression of this gene, however, is inhibited about 5-fold by overproduction of GAL4 when binding sites for either GCN4 (UAS<sub>H</sub>)<sup>18</sup> or HAP1 and HAP2, 3 (UAS<sub>C</sub>)<sup>19</sup> replace the GAL4 binding sites (UAS<sub>G</sub>). GAL4 produced at these high levels similarly inhibits expression of a CYC1:*lacZ* fusion gene about 5-fold. Expression of a HIS4:*lacZ* fusion gene is inhibited about 2-fold by high levels of GAL4 (see Table 1). Cells expressing high levels of some of the GAL4 derivatives, such as GAL4(1–147+768–881), grow somewhat slowly, consistent with

the idea that many genes other than those described here are inhibited by high levels of GAL4.

High levels of GAL4 inhibit inducible transcription from the HIS3 gene, but do not affect constitutive transcription. Struhl has shown that the HIS3 gene is controlled by two promoters: one directs constitutive synthesis of a transcript initiating at +1, the other directs synthesis of a transcript initiating at +12 that is inducible by activators such as GAL4 and GCN4<sup>14,20</sup>. As shown in Fig. 1, transcription from the inducible start site (+12) of the HIS3 promoter is inhibited by high levels of GAL4, while transcription from the constitutive start site (+1) is unaffected. This result is consistent with the idea that squelching results from titration of a specific transcription factor required for only the inducible HIS3 transcript. A TATA-binding protein, which Struhl proposed to distinguish between the two transcripts<sup>8,20,21</sup>, is a plausible candidate for the limiting factor.

The strength of the activating region of a number of GAL4 derivatives, as measured on a gene bearing UAS<sub>G</sub>, correlates well, if not perfectly, with the inhibitory effect measured on a gene lacking UAS<sub>G</sub> (Table 2a and b). Thus, GAL4(1–238), which bears only activating region I, inhibits to a lesser extent than does, for example, the stronger activator GAL4(1–147+768–881), which bears activating region II. When the 33 carboxy-terminal amino acids are removed from GAL4(1–147+768–881), a strong activator and inhibitor, to create GAL4(1–147+768–848), both activities are lost. A short peptide designed to form an amphipathic  $\alpha$ -helix activates and inhibits when fused to GAL4(1–147); a scrambled version of this peptide, with the same amino acid composition, does neither. The behaviour of mutant forms of GAL4(1–238) is also striking (Table 2b): the introduction of 2 or 4 amino acid changes, each of which increases the negative charge of the activating region, makes GAL4(1–238) a more potent activator and a more potent inhibitor. Mutations that remove single acidic residues in the

**Table 1** High levels of GAL4 inhibit gene expression

|                                       | No GAL4 | High levels of GAL4 |
|---------------------------------------|---------|---------------------|
| UAS <sub>G</sub> ---GAL1: <i>lacZ</i> | <1      | 1,415               |
| UAS <sub>H</sub> ---GAL1: <i>lacZ</i> | 200     | 40                  |
| UAS <sub>C</sub> ---GAL1: <i>lacZ</i> | 470     | 115                 |
| UAS <sub>C</sub> ---CYC1: <i>lacZ</i> | 210     | 40                  |
| UAS <sub>H</sub> ---HIS4: <i>lacZ</i> | 345     | 180                 |

A multicopy plasmid encoding the GAL4 gene, expressed from the ADH1 promoter (pGG3), or a control plasmid lacking the GAL4 gene (pGG1), was introduced into *Dga14* yeast bearing the indicated *lacZ* fusion gene and  $\beta$ -galactosidase activity was measured. In this, and subsequent tables,  $\beta$ -galactosidase activity was taken as an indication of promoter activity. The plasmids pGG1 and pGG3 are as described previously<sup>3</sup>. The UAS<sub>G</sub>---GAL1:*lacZ* fusion, in line 1, bears four GAL4 binding sites positioned 310 base pairs (bp) upstream of the start site of transcription (pRY171, ref. 5). The UAS<sub>H</sub>---GAL1:*lacZ* fusion, in line 2, bears the intact HIS4 UAS positioned 128 bp upstream of the transcription start site<sup>3</sup>. The UAS<sub>C</sub>---GAL1:*lacZ* fusion, line 3, bears the intact CYC1 UAS (position –178 to –312 of CYC1, ref. 19) positioned 128 bp upstream of the GAL1 transcription start site. The UAS<sub>C</sub>---CYC1:*lacZ* fusion bears the CYC1 UAS at its normal position –258 from the CYC start (pLG312, ref. 19). The UAS<sub>H</sub>---HIS4:*lacZ* fusion bears the HIS4 UAS 100 bp upstream of the HIS4 transcription start site (ref. 18 and M. Lamphier, unpublished). The HIS4 gene, like the HIS3 gene discussed in the text, is controlled by two promoters, only one of which is inducible by GCN4<sup>27</sup>. We believe this explains why the HIS4 promoter is inhibited only 2-fold by high levels of GAL4. Each *lacZ* fusion gene was integrated in single copy at the *ura3* locus of YM335 (a,  $\Delta$ gal4, *ura3*–52, *ade2*–101, *lys1*, *met*, *his3*). Plasmids were transformed into yeast cells treated with lithium acetate<sup>28</sup> and grown in minimal media lacking histidine and containing 2% galactose, 3% glycerol, and 2% lactic acid, pH6, as carbon sources. Exponentially growing cells were assayed for  $\beta$ -galactosidase activity by standard procedures<sup>4</sup>; the standard error for cultures assayed in triplicate was less than 20%.



**Table 2** Squelching by various GAL4 derivatives correlates with activation

| <i>a</i>             | UAS <sub>H</sub> ---GAL1: <i>lacZ</i> |       | <i>b</i>                 | UAS <sub>H</sub> ---GAL1: <i>lacZ</i> |       |
|----------------------|---------------------------------------|-------|--------------------------|---------------------------------------|-------|
|                      | UAS <sub>G</sub> ---GAL1: <i>lacZ</i> |       |                          | UAS <sub>G</sub> ---GAL1: <i>lacZ</i> |       |
| No GAL4              | 200                                   | <1    | No GAL4                  | 200                                   | <1    |
| GAL4                 | 40                                    | 1,415 | GAL4                     | 43                                    | 3,320 |
| GAL4 (1-147+768-881) | 36                                    | 624   | GAL4 (1-238)             | 126                                   | 710   |
| GAL4 (1-147)+B17     | 49                                    | 398   | GAL4 (1-238)             |                                       |       |
| GAL4 (1-238)         | 88                                    | 130   | Arg <sub>166</sub> → Thr | 52                                    | 1,910 |
| GAL4 (1-147)+AH      | 52                                    | 121   | Lys <sub>188</sub> → Gln |                                       |       |
| GAL4 (79-881)        | 66                                    | <1    | Arg <sub>166</sub> → Trp | 59                                    | 2,565 |
| GAL4 (1-147)         | 200                                   | <1    | Lys <sub>188</sub> → Gln |                                       |       |
| GAL4 (1-147)+SH      | 200                                   | <1    | Arg <sub>207</sub> → Pro |                                       |       |
| GAL4 (1-147+768-848) | 158                                   | <1    | Arg <sub>231</sub> → Ser |                                       |       |
|                      |                                       |       | Asp <sub>149</sub> → His | 207                                   | 110   |
|                      |                                       |       | Asp <sub>199</sub> → Tyr | 195                                   | 17    |

Plasmids encoding the indicated GAL4 derivatives, each expressed from the ADH1 promoter, were introduced into  $\Delta$ gal4 yeast bearing a GAL1:*lacZ* fusion with either the HIS4 UAS or the GAL UAS upstream and  $\beta$ -galactosidase activity was measured. Note that different yeast strains were used in (*a*) and (*b*) and multicopy plasmids were used in (*a*), whereas single copy plasmids were used in (*b*). *a*, Plasmids encoding no GAL4 (pGG1) and intact GAL4 (pGG3) are as described<sup>3</sup>. Deletion derivatives of GAL4 bearing the indicated amino acids GAL4 (1-147), (1-238), (1-147+768-881) and (1-147+768-848) are encoded by the plasmids pMA241, pMA246, pMA236, and pMA238 respectively<sup>9</sup>. GAL4(1-147)+AH, encoded by pEG50, and GAL4(1-147)+SH, encoded by pEG52, each bear a 15-amino acid peptide with a net charge of -3, the first (AH) was designed to form an amphipathic  $\alpha$ -helix<sup>13</sup>. GAL4(1-147)+B17 bears an *E. coli* encoded activating region whose isolation and sequence is described<sup>12</sup>. The plasmid encoding GAL4(79-881) (pGG19) was constructed by restoring GAL4 carboxy-terminal sequences of a GAL4(79-147)+ $\beta$ -galactosidase fusion described by Silver and colleagues<sup>30</sup>. All strains and methods are as described in Table 1. *b*, Single copy plasmids encoding the indicated GAL4 derivatives expressed from the ADH1 promoter were introduced into GGY1 (ura3-52, leu2-2,112, his3 $\Delta$ 200,  $\Delta$ gal4,  $\Delta$ gal180) (ref. 11) bearing the *lacZ* fusions described in Table 1 integrated in single copy at the ura3 locus. As can be seen by comparing *a* and *b*, the activity of any given GAL4 derivative varies from strain to strain. Plasmids encoding no GAL4 (pBRHAC), intact GAL4 (pGG22), and GAL4(1-238) (pGG23) are as described<sup>11</sup>. Isolation and characterization of the mutants in activating region I has been described, except the quadruple mutant which results from *in vitro* recombination of four previously described mutants<sup>11</sup>. Note that the effects of mutations which increase the activity of region I as an activator are roughly additive, so that the quadruple mutant is 80% as active as intact GAL4. All methods are as described in Table 1.

activating region make GAL4(1-238) a weaker activator which is unable to inhibit expression of genes lacking GAL4 binding sites.

Squelching is not caused by binding of GAL4 to DNA, for example to cryptic sites in the coding region. Table 2*a* shows that the derivative GAL4(79-881), which does not bind DNA, and therefore cannot activate although the activating regions are present<sup>3,22</sup>, inhibits. GAL4(1-147) which binds DNA *in vivo* but does not activate<sup>3,9</sup> does not inhibit.

High levels of expression of GAL80, a negative regulator of activation by GAL4, reduces squelching. Table 3 describes the behaviour of a strain in which a GAL4 gene expressed from the strong ADH1 promoter has been integrated in the chromosome. Expression of a GAL1:*lacZ* fusion gene bearing UAS<sub>H</sub> is stimulated in this strain by the introduction of a multicopy plasmid expressing the negative regulator GAL80. Previous experiments have shown that GAL80 interacts with an essential part of activating region II of GAL4, and the GAL4/GAL80 complex, although it binds DNA, does not activate transcription<sup>15-17,23</sup>. We interpret the stimulatory effect of GAL80 in this experiment as a consequence of GAL80 masking the activating region of GAL4, thereby relieving the squelching.

We do not think it is likely that our results can be explained by the formation of mixed oligomers between the GAL4 derivatives and each of the activators (GCN4, HAP1, HAP2, 3) of our test genes. GAL4(1-147), for example, forms stable dimers *in vitro* (M. Carey, H. Kakidani and M.P., manuscript in preparation) and efficiently binds UAS<sub>G</sub> *in vivo*, but does not squelch.

We imagine that the interaction between an activating region and a transcription factor, the TATA-binding protein for example, although facilitated by DNA binding, can also occur free in solution. At high activator concentrations, this reaction would sequester the factor and inhibit transcription. The effect would be first noticed with genes lacking the DNA binding sites for the activator, although at very high activator concentrations even genes with the binding sites would begin to be inhibited, a result we have observed (data not shown). We believe our results parallel those of Triezenberg and colleagues who find that VP16, the very strong transcriptional activator found in

**Table 3** High levels of GAL80 relieve squelching

|          | UAS <sub>H</sub> ---GAL1: <i>lacZ</i> |
|----------|---------------------------------------|
| no GAL80 | 7                                     |
| GAL80    | 20                                    |

A multicopy plasmid encoding the GAL80 gene (Yep213-80) or a control plasmid lacking this gene (Yep213) was introduced into a yeast strain bearing GAL4 expressed from the ADH1 promoter integrated at the gal4 locus, and a GAL1:*lacZ* fusion gene with the HIS4 UAS upstream integrated at the ura3 locus, and the  $\beta$ -galactosidase activity was measured. The parent yeast strain used was GGY2 ( $\alpha$ , ura3-52, leu2-2, 112, lys2,  $\Delta$ gal4, his3). There was no difference in the  $\beta$ -galactosidase activity produced by the UAS<sub>H</sub>---GAL1:*lacZ* fusion gene when multicopy plasmids bearing or lacking the GAL80 gene were introduced into a  $\Delta$ gal4 strain.  $\beta$ -galactosidase activity from a GAL1:*lacZ* fusion gene with the GAL UAS upstream is reduced about 4-fold when GAL80 is expressed at high levels, even in galactose-containing media (data not shown and ref. 17). All methods as described in Table 1.

herpes simplex virus, inhibits transcription of genes lacking the appropriate binding sites<sup>24,25</sup>. There is also a striking correlation between activation and inhibition in this case: derivatives of VP16 that activate efficiently also inhibit efficiently, and deletion derivatives impaired as activators are also impaired as inhibitors. Similarly, GAL4(1-147+VP16), a hybrid protein containing the DNA-binding region of GAL4 fused to the highly acidic carboxy-terminus of VP16, activates transcription of genes bearing UAS<sub>G</sub> very efficiently, but inhibits expression of other genes both in mammalian cells and yeast<sup>26</sup>. As the degree of squelching should depend on both the concentration of the activator and the strength of the activating region, there may be a limit to the possible strength and/or concentration of any given activator in viable eukaryotic cells.

We thank Howard Himmelfarb, Anita Meiklejohn and Ivan Sadowski for helpful discussions, Asa Abeliovitch, Ed Giniger, Marc Lamphier and Jun Ma for plasmids and strains, and Mike Carey for advice on primer extensions. We acknowledge the support of the National Institutes of Health.

Received 25 May; accepted 11 July 1988.

1. Bram, R. & Kornberg, R. D. *Proc. natn. Acad. Sci. U.S.A.* **82**, 43-47 (1985).
2. Giniger, E., Varnum, S. M. & Ptashne, M. *Cell* **40**, 767-774 (1985).
3. Keegan, L., Gill, G. & Ptashne, M. *Science* **231**, 699-704 (1986).
4. West, R. Jr, Yocum, R. & Ptashne, M. *Molec. cell. Biol.* **4**, 2467-2478 (1984).
5. Yocum, R., Hanley, S., West, R. Jr & Ptashne, M. *Molec. cell. Biol.* **4**, 1985-1988 (1984).
6. Brent, R. & Ptashne, M. *Cell* **43**, 729-736 (1985).
7. Ptashne, M. *Nature* **322**, 697-701 (1986).
8. Struhl, K. *Cell* **49**, 295-297 (1987).
9. Ma, J. & Ptashne, M. *Cell* **48**, 847-853 (1987).
10. Hope, I. A. & Struhl, K. *Cell* **46**, 885-894 (1986).
11. Gill, G. & Ptashne, M. *Cell* **51**, 121-126 (1987).
12. Ma, J. & Ptashne, M. *Cell* **51**, 113-119 (1987).
13. Giniger, E. & Ptashne, M. *Nature* **330**, 670-672 (1987).
14. Struhl, K. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7865-7869 (1984).
15. Ma, J. & Ptashne, M. *Cell* **50**, 137-142 (1987).
16. Johnston, S. A., Salmeron, J. M. & Dincher, S. S. *Cell* **50**, 143-146 (1987).
17. Nogi, Y., Shimada, H., Matsuzaki, Y., Hashimoto, H. & Fukusawa, T. *Molec. Gen. Genet.* **195**, 29-34 (1984).
18. Silverman, S. J., Rose, M., Botstein, D. & Fink, G. R. *Molec. cell. Biol.* **2**, 1212-1219 (1982).
19. Guarente, L. & Mason, T. *Cell* **32**, 1279-1286 (1983).
20. Struhl, K. *Molec. cell. Biol.* **6**, 3847-3853 (1986).
21. Chen, W. & Struhl, K. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2691-2695 (1988).
22. Johnston, S. A., Zavortink, M. J., Debouck, C. & Hopper, J. E. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6553-6558 (1986).
23. Lue, N. F., Chasman, D. I., Buchman, A. R. & Kornberg, R. D. *Molec. cell. Biol.* **7**, 3446-3451 (1987).
24. Triezenberg, S. J., Kingsbury, R. C. & McKnight, S. L. *Genes Dev.* **2**, 718-729 (1988).
25. Triezenberg, S. J., LaMarco, K. L. & McKnight, S. L. *Genes Dev.* **2**, 730-742 (1988).
26. Sadowski, I., Ma, J. & Ptashne, M. *Nature*, submitted.
27. Nagawa, F. & Fink, G. R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8557-8561 (1985).
28. Ito, H., Fukuda, Y., Murata, K. & Kimura, A. *J. Bact.* **53**, 163-168 (1983).
29. Struhl, K. *Nucleic Acids Res.* **13**, 8587-8601 (1985).
30. Silver, P. A., Keegan, L. P. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5951-5955 (1984).

## Antisense RNA inhibition of polygalacturonase gene expression in transgenic tomatoes

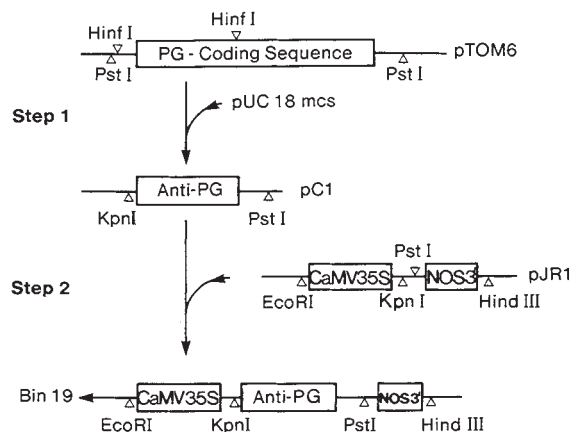
C. J. S. Smith\*, C. F. Watson\*, J. Ray†, C. R. Bird†, P. C. Morris\*, W. Schuch†‡ & D. Grierson\*‡

\* Department of Physiology and Environmental Science, University of Nottingham, Faculty of Agricultural Science, Sutton Bonington, Loughborough LE12 5RD, UK

† ICI Seeds, Plant Biotechnology Section, PO Box 11, The Heath, Runcorn, Cheshire WA7 4QE, UK

Regulation of expression of specific genes by antisense RNA is a naturally occurring mechanism in bacteria<sup>1,2</sup>, although gene regulation by this mechanism has not yet been observed in higher eukaryotes. However, antisense RNA has been shown to reduce expression of specific genes when injected into frog oocytes<sup>3</sup> and *Drosophila* embryos<sup>4</sup>. Inhibition of expression of artificially introduced genes has been demonstrated by transient expression of antisense RNA constructs in mammalian cells<sup>5,6</sup>, and plant protoplasts<sup>7</sup>, and by stable expression in transgenic plants<sup>8</sup>. Here, we report a striking inhibition of expression of the endogenous, developmentally regulated gene for polygalacturonase in stably transformed tomato expressing antisense RNA.

Polygalacturonase (PG) plays an important role in fruit softening by partially solubilizing the pectin fraction of the cell walls<sup>9-11</sup>. It is synthesized *de novo*<sup>12</sup> as a result of the accumulation of PG mRNA<sup>13-15</sup> when fruit ripen. We have previously isolated and characterized a complementary DNA (cDNA) clone (pTOM 6)<sup>16</sup> from tomato (*Lycopersicon esculentum* Mill. cv Ailsa Craig) and also genomic clones for the PG gene<sup>17</sup>. Only one PG gene has been detected<sup>17</sup>; it is therefore possible that the three structurally related isoenzyme forms found in ripe fruit<sup>9,18-20</sup> arise from post-translational modification of the same polypeptide<sup>21-23</sup>. We have developed a homologous transformation system to study the control of expression of ripening genes<sup>17</sup> and have used this to introduce a construct designed to express PG antisense RNA constitutively into tomato plants. A 730-base-



**Fig. 1** Construction of pJR16A. Step 1: A 730 bp *HindIII* fragment from the 5' end of the PG cDNA clone pTOM 6<sup>16</sup> was isolated and the ends filled in with Klenow polymerase. The fragment was cloned into the *SmaI* site of a pUC18 multiple cloning site. Step 2: The *KpnI*-*PstI* fragment from a clone (pC1) containing the cDNA fragment in the appropriate orientation was ligated into the same sites of the pUC based expression vector pJR1 (J.R., unpublished). This construct contains a 528-bp CaMV 35S promoter fragment (nucleotides 6,909-7,437, ref. 27) in which a transcription start site has been identified at position 7,435 (ref. 24). The antisense gene was excised by a partial *EcoRI*-*HindIII* digest and inserted into *EcoRI*-*HindIII* digested Bin 19 (ref. 26) to give pJR16A. The construction of pJR16A was confirmed by DNA sequence analysis. The vector was transferred from *E. coli* TG-2 to *A. tumefaciens* LBA4404 by triparental mating with pRK2013 in *E. coli* HB101 (ref. 26). Restriction analysis of plasmid DNA from *A. tumefaciens* confirmed that the construct was intact.

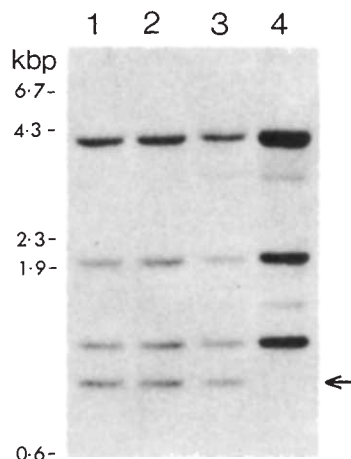
pair (bp) *HindIII* fragment from the 5' end of the PG cDNA, including a 50-bp untranslated region and the translation start site, was fused, in the inverted orientation, to the cauliflower mosaic virus (CaMV) 35S RNA promoter<sup>24</sup> and the 3' end of the nopaline synthase (*nos*) gene<sup>25</sup>. This hybrid gene was introduced into the *Agrobacterium tumefaciens* binary transformation vector Bin 19<sup>26</sup>. This plasmid, pJR16A (Fig. 1) was transferred from *Escherichia coli* strain TG-2 to *A. tumefaciens* strain LBA 4404 by triparental mating and used to transform tomato stem segments<sup>17</sup>. Transformed plants were selected by their ability to grow and form roots on media containing kanamycin, and grown to maturity. Three of these plants, found to express antisense RNA in leaves (GR15, GR16 and GR17), were selected for further analysis.

Genomic Southern blots of leaf DNA from the three transformed plants were probed with radiolabelled pTOM 6 insert. A 1,230-bp hybridizing fragment was present in transformed but not in untransformed leaf DNA (Fig. 2). This fragment is the CaMV promoter and part of the PG antisense sequence from pJR16A (see Fig. 1). The relative intensities of the antisense and endogenous PG bands indicate that the copy number of the antisense gene is similar to that of PG (believed to be one<sup>17</sup>).

Northern analysis of poly(A)<sup>+</sup> RNA from leaves and fruit of GR16 was used to show that the transformed plants expressed the PG antisense gene. The membranes were hybridized with strand-specific probes to differentiate between transcripts from the antisense and the endogenous PG gene. Two major transcripts (960 and 660 nucleotides) in GR16 fruit hybridized to the probe specific for the antisense strand (Fig. 3a, lanes 5, 6). The largest corresponds to the predicted size of the full-length nonadenylated transcripts (917-929 nucleotides) of the PG antisense gene from pJR16A. In leaf RNA we also observed a third, smaller band of 520 nucleotides (data not shown). These smaller bands probably result from premature termination of transcription. There are potential polyadenylation-like sequences in the antisense sequence which could give rise to partial transcripts

‡ To whom correspondence should be addressed.





**Fig. 2** Genomic Southern of antisense tomato plants. Approximately 2 µg genomic DNA digested with *EcoRI* and *HindIII* was probed with the *PstI* insert of pTOM 6. Lane 1, GR15; lane 2, GR16; lane 3, GR17; lane 4, untransformed plant. The 1,230-bp hybridizing fragment (arrow) in lanes 1, 2 and 3 contains the CaMV 35S promoter and part of the antisense gene. This fragment is not found in DNA from untransformed tomato (lane 4).

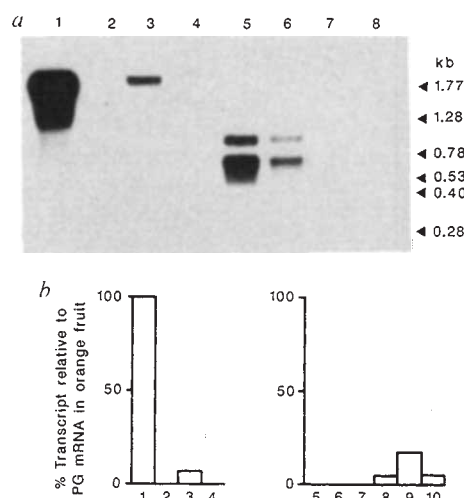
**Methods.** Genomic DNA was extracted from frozen leaves as described<sup>28</sup>. Approximately 2 µg of genomic DNA was digested with *EcoRI* and *HindIII*, separated on a 0.8% agarose gel and blotted as described<sup>29</sup>. The membranes were prehybridized in 5 × SSPE, 0.75 M NaCl, 50 mM NaH<sub>2</sub>PO<sub>4</sub> · H<sub>2</sub>O, 5 mM EDTA, 1% SDS, 100 µg ml<sup>-1</sup> denatured herring sperm DNA at 65 °C for 4 h and hybridized with oligo-labelled pTOM 6 insert (Pharmacia Labelling kit) in the same buffer for 18 h at 65 °C. The membranes were washed in 0.1% SSC, 1% SDS at 65 °C and autoradiographed.

of these sizes. In untransformed tissue no antisense transcripts were detected (Fig. 3a, lanes 7, 8). PG mRNAs of identical size were detected in ripe fruit of untransformed and antisense plants (Fig. 3a, lanes 1, 3). However, there was a striking reduction in the amount of PG mRNA in ripe fruit from GR16. No PG mRNA was detected in green fruit.

Quantitative RNA dot hybridization analysis showed that GR16 expressed PG antisense RNA in ripe and unripe fruit and leaves at similar concentrations (Fig. 3b). As fruit ripen, the endogenous PG mRNA is normally at its peak concentration in orange fruit<sup>13,14</sup>. At this stage, PG mRNA in GR16 was 6% of that in untransformed fruit. Also in GR16, the level of antisense RNA was 60% of that of PG mRNA.

Fractionation of cell-wall-associated proteins by SDS-polyacrylamide gel electrophoresis (PAGE) showed that there was a substantial accumulation of PG protein in ripe control fruit, but much less in fruit expressing the antisense gene (Fig. 4a). PG enzyme activity was also found to be significantly reduced in ripening fruit from each of the three separate transformants. The greatest reduction was in GR16 where 10% of the normal PG activity was detected (Fig. 4a), which correlated well with the reduction in PG protein. Two bands representing cell-wall-associated proteins of relative molecular mass 25,000 (25K) and 30K seem to have increased in GR15 and GR16. The explanation for this is unclear.

PG is usually accumulated in parallel with the pigment lycopene during ripening. As expected, fruit expressing antisense PG RNA synthesized similar amounts of lycopene as control fruit. This allowed us to compare the levels of extractable PG activity in fruit expressing normal and antisense PG RNA at different ripening stages. This analysis showed that PG activity was very substantially reduced in GR16 fruit at all stages of ripening, even in fruit that had developed full colour (Fig. 4b). A major reduction in PG activity would be expected to reduce depolymerization of cell-wall pectin during ripening, leading to

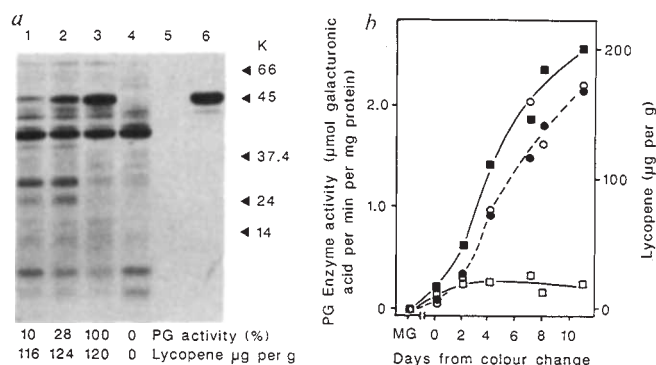


**Fig. 3** a, Northern blot of poly(A)<sup>+</sup> RNA from untransformed and transformed fruit. RNA (1 µg per lane) from untransformed and GR16 fruit was hybridized to a PG sense-strand-specific probe (lanes 1-4) and an antisense-strand-specific probe (lanes 5-8). Orange untransformed fruit (lanes 1, 8); green untransformed fruit (lanes 2, 7); orange transformed fruit (lanes 3, 6); green transformed fruit (lanes 4, 5). b, Histograms of the dot blot analysis of poly(A)<sup>+</sup> mRNA from untransformed and transformed (GR16) plants hybridized to a sense-strand-specific probe (1-4) and an antisense-strand-specific probe (5-10). Orange untransformed fruit (1, 5); green untransformed fruit (2, 6); orange transformed fruit (3, 8); green transformed fruit (4, 9); untransformed leaf (7) and transformed leaf (10). Results are expressed as percentage of PG mRNA level in orange untransformed fruit.

**Methods.** RNA was extracted from tomato pericarp<sup>30</sup> and leaves<sup>31</sup>. Polyadenylated RNA was isolated by oligo(dT) cellulose chromatography. Glyoxal-denatured poly(A)<sup>+</sup> RNA was either fractionated on a 1.2% agarose gel before blotting onto nylon membrane (Hybond N, Amersham), or dotted directly onto nylon membrane. Hybridization was performed at 65 °C in 10 × NTE, 0.3 M NaCl, 2 mM EDTA, 60 mM Tris-HCl pH 8.0 1% SDS, 200 µg ml<sup>-1</sup> salmon sperm DNA, 5 × Denhardt's solution. Membranes were washed at 65 °C in 0.2 × NTE, 0.1% SDS. PG sense and antisense-specific probes were synthesized using a TransProbe T kit (Pharmacia). The *KpnI*-*PstI* fragment from a clone (C4) containing the cDNA fragment in the opposite orientation to clone C1 (see Fig. 1) was ligated into pBSM13<sup>+</sup> (Stratagene) and cut with *KpnI* and *PstI* to form pBS-6. Antisense-strand-specific probe was transcribed by T7 polymerase from *PstI* cut pBS-6, sense-strand-specific probe by T3 polymerase from *EcoRI* cut pBS-6.

decreased softening<sup>33</sup>. This might improve fruit shelf-life and processing quality. Attempts to measure softening by compressibility show no differences between GR16 and control fruit. However, the physiological and biochemical changes that cause softening are complex and may not relate only to the pectin fraction. Further work is in progress to define softening and the role of PG more precisely by measurements of the amount and degree of pectin depolymerization during ripening of antisense and normal fruit.

The antisense PG gene described in this paper is transcribed constitutively under the direction of the CaMV 35S RNA promoter in leaves and fruit of transformed plants. Analysis of 12 first generation (F<sub>1</sub>) plants of GR16 indicates that it is stably inherited. PG activity was substantially reduced in ripe fruit of eight F<sub>1</sub> plants, but was normal in the remaining four. Genomic Southern analysis has shown that the gene was present in F<sub>1</sub> plants with reduced PG activity but was absent in those with normal activity (data not shown). Therefore, reduced PG activity segregates with the antisense gene in the F<sub>1</sub> generation. The PG antisense RNA is stable and is polyadenylated and there are at least three transcripts, one of which seems full-length. We do not know which of these transcripts is the most significant in



**Fig. 4** PG expression during ripening of transformed and untransformed fruit. *a*, Cell-wall-associated proteins separated by SDS-PAGE. Lane 1, ripe GR16; lane 2, ripe GR15; lane 3, ripe control; lane 4, mature green control; lane 5, blank; lane 6, purified PG2. Samples 1–3 were extracted seven days after the start of ripening. PG activity and lycopene content of the fruit are shown. Control fruit (lanes 3, 4) were from plants that had been through the transformation procedure but without antisense gene constructs. *b*, PG activity of cell-wall-associated proteins and lycopene content during ripening. PG activity in untransformed (■) or GR16 (□) fruit. Lycopene content of untransformed (●), or GR16 (○) fruit. MG samples were from mature green fruit.

**Methods.** PG was extracted and assayed as described previously<sup>9</sup>. Protein was measured with Pierce protein assay reagent (Pierce Chemical). Proteins were fractionated by electrophoresis through a denaturing 15% SDS-urea polyacrylamide gel and stained with Coomassie brilliant blue dye. Pigment was extracted from pericarp by vigorously homogenizing in a 6:4 mixture of hexane and acetone and scanned in a spectrophotometer. 1 mg ml<sup>-1</sup> lycopene gives an absorbance of 320 at 502 nm<sup>32</sup>.

the inhibition of PG synthesis during ripening. The effect of the antisense construct on PG levels is variable in different transformants. This may be related to differences in the steady-state levels of antisense RNA generated due to positional effects associated with particular DNA insertion events.

Our results indicate that PG antisense RNA generated in transgenic tomato plants causes a reduction in PG enzyme activity when the fruit ripen. In GR16 the effect of the antisense RNA is dramatic, resulting in a reduction of 90% in the expected level of PG activity in ripe fruit. The precise mechanism of inhibition of enzyme synthesis is as yet unknown. The low level of PG mRNA observed in ripe fruit of GR16 may point to some process occurring in the nucleus, interference with transcription, processing or transport, for example, rather than inhibition of translation by the formation of RNA/RNA hybrids in the cytoplasm. Alternatively, the low levels of PG mRNA could be caused by the selective degradation of double-stranded RNA hybrids as Rothstein *et al.* suggest<sup>8</sup>. Further work is required before we can distinguish between these possibilities. However the successful down-regulation of expression of an endogenous plant gene by antisense RNA demonstrates that this approach is a practical general strategy for studies on development, gene expression and plant breeding.

We thank A. Bird, N. Halstead and S. Hughes for technical assistance and Dr G. A. Tucker for purified PG. This work was supported by a grant to D.G. from the SERC Biotechnology Directorate and carried out under MAFF licences.

**Note added in proof:** While this paper was being reviewed we learnt that the down-regulation endogenous plant gene by antisense RNA has been observed by van der Krol *et al.*<sup>34</sup>.

5. Izant, J. G. & Weintraub, H. *Cell* **36**, 1007–1015 (1984).
6. Kim, S. K. & Wold, B. J. *Cell* **42**, 129–138 (1985).
7. Ecker, J. R. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5372–5376 (1986).
8. Rothstein, S. J., DiMaio, J., Strand, M. & Rice, D. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8439–8443 (1987).
9. Tucker, G. A., Robertson, N. E. & Grierson, D. *Eur. J. Biochem.* **112**, 119–124 (1980).
10. Themmen, A. P. N., Tucker, G. A. & Grierson, D. *Plant Physiol.* **69**, 122–124 (1982).
11. Crookes, P. R. & Grierson, D. *Pl. Physiol.* **72**, 1088–1093 (1983).
12. Tucker, G. A. & Grierson, D. *Planta* **155**, 64–67 (1982).
13. Grierson, D. *et al. Phil. Trans. R. Soc. B* 399–410 (1986).
14. Dellapenna, D., Alexander, D. C. & Bennett, A. B. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6420–6424 (1986).
15. Biggs, M. S., Harriman, R. W. & Handa, A. K. *Pl. Physiol.* **81**, 395–403 (1986).
16. Grierson, D. *et al. Nucleic Acids Res.* **14**, 8595–8603 (1986).
17. Bird, C. R. *et al. Pl. molec. Biol.* (submitted).
18. Tucker, G. A., Robertson, N. G. & Grierson, D. *Eur. J. Biochem.* **115**, 87–90 (1981).
19. Ali, Z. M. & Brady, C. J. *Aust. J. Pl. Physiol.* **9**, 155–169 (1982).
20. Pressey, R. *HortScience* **21**, 1183–1185 (1986).
21. Grierson, D. *Oxford Surveys Pl. molec. cell Biol.* **3**, 364–383 (1986).
22. Dellapenna, D. & Bennett, A. B. *Pl. Physiol.* (in the press, 1988).
23. Pressey, R. *Eur. J. Biochem.* **144**, 217–221 (1984).
24. Guilley, H., Dudley, R. K., Jonard, G., Balazs, E. & Richards, K. E. *Cell* **30**, 763–773 (1982).
25. Depicker, A., Stachel, S., Dhaese, P., Zambryski, P. & Goodman, H. M. *J. molec. appl. Genet.* **1**, 561–573 (1982).
26. Bevan, M. W. *Nucleic Acids Res.* **12**, 8711–8721 (1984).
27. Franck, G., Guille, M., Jonard, G., Richards, K. & Hirth, L. *Cell* **21**, 285–294 (1980).
28. Raeder, U. & Broda, P. *Lett. appl. Microbiol.* **1**, 17–20 (1985).
29. Chomczynski, P. & Qasbi, P. K. *Biochem. biophys. Res. Commun.* **122**, 340–344 (1984).
30. Grierson, D., Slater, A., Speirs, J. & Tucker, G. A. *Planta* **163**, 263–277 (1985).
31. Smith, C. J. S., Slater, A. & Grierson, D. *Planta* **168**, 94–100 (1986).
32. Tomes, M. L. *Bot. Gaz.* **124**, 180–185 (1963).
33. Seymour, G. B., Harding, S. E., Taylor, A. J., Hobson, G. E. & Tucker, G. A. *Phytochemistry* **26**, 1871–1875 (1987).
34. van der Krol, A. R. *et al. Nature* **333**, 866–869 (1988).

## Quantum simulation of ferrocyclochrome c

Chong Zheng\*, Chung F. Wong\*,  
J. Andrew McCammon\* & Peter G. Wolynes†

\* Department of Chemistry, University of Houston, Houston, Texas 77004, USA

† Noyes Laboratory, University of Illinois, Urbana, Illinois 61801, USA

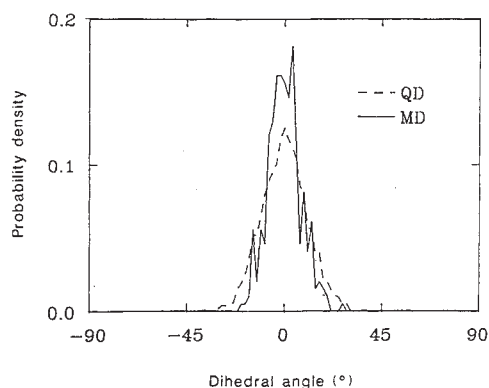
The dramatic progress in the understanding of the dynamics of biomolecules has been largely fuelled by computer simulations based on the law of classical mechanics<sup>1</sup>. However in some respects biomolecules are at the borders of the domain of applicability of classical mechanics. The role of quantum mechanical effects in biomolecular structure and function is therefore worth investigating. Here we present preliminary results from a quantum simulation of a protein and contrast them with results from full classical simulations. The most significant differences are found in motions of high frequency, such as bond stretching or the torsional oscillation of groups that bear hydrogen atoms. The amplitudes of such motions are significantly increased by the penetration of atoms into classically forbidden regions. These differences will directly influence the rates of such processes as proton and electron transfer.

Experimentally we know there are several phenomena in which quantum effects are relevant to biomolecular dynamics. Nuclear tunnelling effects have been observed in the binding reactions of a haem proteins<sup>2</sup> and in electron transfer reactions<sup>3–5</sup>. The fact that some biological electron transfer reactions become temperature independent at temperatures lower than ~150 K suggests that there are substantial quantum effects even under physiological conditions<sup>3,6</sup>. Large isotope effects indicate an important role for quantum effects in proton transfer reactions<sup>7</sup>. Tunnelling between conformational substates may contribute to the thermal properties of proteins at low temperatures<sup>8</sup>. Thermal factors, measured by X-ray diffraction and the Mössbauer effect are influenced by quantum effects, especially at low temperature<sup>9,10</sup>. Finally because many modes of protein motion are high in frequency, zero-point motion effects contribute to overall thermodynamics<sup>11</sup>.

Received 2 May; accepted 23 July 1988.

1. Minzuno, T., Chou, M. & Inouye, M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1966–1970 (1984).
2. Simons, R. W. & Kleckner, N. *Cell* **34**, 683–691 (1983).
3. Melton, D. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 144–148 (1985).
4. Rosenberg, U. B., Preiss, A., Seifert, E., Jackie, H. & Knipple, D. C. *Nature* **313**, 703–706 (1985).





**Fig. 1** Asn 52 NH<sub>2</sub> dihedral angle distributions from the classical (—) and quantum (---) simulations. The distribution is of the deviations from the average dihedral angle in the fluctuating protein.

The method we used is based on the well known isomorphism between the Feynman path integral formulation of quantum statistical mechanics<sup>12</sup> and the classical statistical mechanics of polyatomic molecules which can be investigated with molecular dynamics (MD) methods<sup>13</sup>. The quantum canonical partition function for a single particle can be written as

$$Q(\beta) = \text{Tr} e^{-\beta H} = \int dx_1 \langle x_1 | e^{-\beta H} | x_1 \rangle$$

$$= \int dx_1 dx_2 \dots dx_P \langle x_1 | e^{-(\beta/P)H} | x_2 \rangle$$

$$\times \langle x_2 | e^{-(\beta/P)H} | x_3 \rangle \dots \langle x_P | e^{-(\beta/P)H} | x_1 \rangle \quad (1)$$

by inserting identities  $\int dx_n |x_n\rangle \langle x_n|$ . For very large  $P$ ,  $(\beta/P)H$  is small and  $\langle x_n | e^{-(\beta/P)H} | x_{n+1} \rangle$  is the imaginary time propagator of a particle subject to a very weak potential<sup>12-14</sup>:

$$\langle x_n | e^{-(\beta/P)H} | x_{n+1} \rangle \approx \left( \frac{mP}{2\pi\beta\hbar^2} \right)^{3/2}$$

$$\times \exp \left\{ -\frac{mP(x_n - x_{n+1})^2}{2\beta\hbar^2} - \frac{\beta V(x_n)}{P} \right\} \quad (2)$$

where  $\beta = 1/(k_B T)$ ,  $T$  is the temperature of the system,  $k_B$  and  $\hbar$  are the Boltzmann's and Planck's constants, respectively. Replacing the propagators in (1) with (2), we can write the quantum partition function as

$$Q(\beta) = \lim_{P \rightarrow \infty} Q_P(\beta)$$

$$\equiv \lim_{P \rightarrow \infty} \left( \frac{mP}{2\pi\beta\hbar^2} \right)^{3P/2} \int \dots \int dx_1 \dots dx_P$$

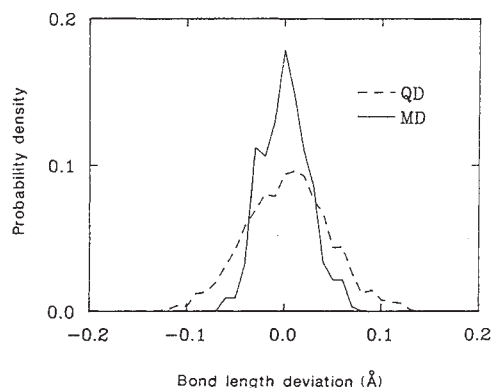
$$\times \exp \{ -\beta \Phi_P(x_1 \dots x_P; \beta) \} \quad (3)$$

where  $x_{P+1} = x_1$ , and

$$\Phi_P(x_1 \dots x_P; \beta) = \left( \frac{1}{2} \right) k_P \sum_{i=1}^P (x_i - x_{i+1})^2 + \frac{1}{P} \sum_{i=1}^P V(x_i) \quad (4)$$

$$k_P = \frac{mP}{\beta^2 \hbar^2}$$

which is equivalent to the classical partition function of  $P$  particles connected sequentially by harmonic springs with force constant  $k_P$ . The sequential connectivity reflects the causality of the quantum propagator and the quadratic force comes from the gaussian-type diffusion property of the density matrix that obeys the Liouville equation. The approximation made in equation (2) is valid when the first term in the exponential



**Fig. 2** Fe-N(His 18) bond length distributions from the classical (—) and quantum (---) simulations. The distribution is of the deviations from the average bond length in the fluctuating protein.

dominates. Thus, if  $\sigma$  is a characteristic length scale on which  $V(x)$  changes, then for  $\beta k_P \sigma^2 \gg 1$  or  $P \gg (\beta \hbar^2)/(m \sigma^2)$ ,  $Q(\beta)$  can be approximated by  $Q_P(\beta)$ . Except for an easily determined multiplicative factor playing no role in the dynamics,  $Q_P$  can also be written as

$$Q_P(\beta) = \int dp_1 \dots dp_P \int dx_1 \dots dx_P \exp [-\beta H_{\text{eff}}] \quad (5)$$

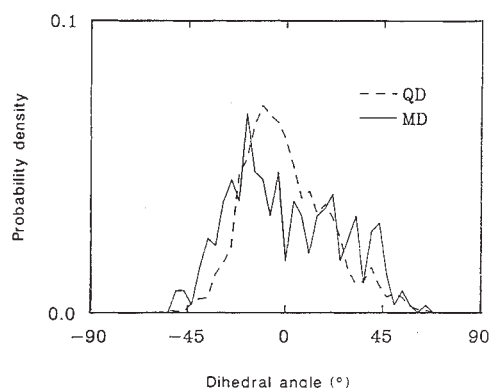
where

$$H_{\text{eff}} = \sum_{i=1}^P \frac{p_i^2}{2m'} + \Phi_P(x_1 \dots x_P; \beta) \quad (6)$$

and  $m'$  is an arbitrary mass chosen to determine the time step of the MD sampling.

In this formalism, each quantum particle is represented by a ring (or 'necklace') of  $P$  classical particles (or 'beads'). The cyclic topology is due to the fact that the quantum partition function in equation (1) is the integral of an imaginary time propagator leaving and returning to the same point in space. Because of the uncertainty principle, the position of a quantum particle cannot be sharply defined at thermal equilibrium. The breadth of the particle's thermal wavepacket is reflected in the average width of the necklace. The bead at position  $i$  is connected to two adjacent beads by a quantum spring with force constant  $k_P$  and is subjected to an external potential  $(1/P)V(x_i)$ . If the potential is from another quantized atom (a necklace), then only the beads in the same position (same  $i$ ) in both necklaces interact with each other<sup>12-14</sup>. Through this isomorphism a quantum system can be simulated by the MD method. This procedure has been used to simulate a wide range of simple systems<sup>14</sup>. One of the powerful attributes of this approach to quantum simulations is that it can deal with complicated anharmonic Hamiltonians. Most treatments of quantum effects on nuclear motion in proteins have been tied to harmonic approximations. Although these approximations may be valid for many of the high frequency modes of motion with large quantum effects, they cannot deal well with tunnelling motions, as in rotations of side chains, and may not do well when quantum harmonic modes are strongly coupled to anharmonic low frequency modes.

We used GROMOS empirical potential parameters<sup>15</sup> in our study. Although the mass of each bead  $m'$  in (6) is chosen to recover the pre-integral factor in (3), in practice it may be selected to give an effective MD sampling over the coordinate space, as a classical ensemble-averaged quantity is independent of the mass. From exploratory simulations on small test systems, we found that good sampling could be achieved for  $m'$  equal to the proton mass for all quantized atoms. The 1.5 Å X-ray crystallographic structure of tuna ferrocycytochrome *c* (cyt *c*),



**Fig. 3** Haem propionate rotational distributions from the classical (—) and quantum (---) simulations. The distribution is of the deviations from the average rotational angle in the fluctuating protein.

described by Takano and Dickerson<sup>16</sup>, was thermalized and equilibrated to 300 K (the details to be described elsewhere). Classical and quantum simulations were carried out in parallel for 10 picoseconds. Residues of interest in electron transfer, His 18, Tyr 48, Thr 49, Asn 52, Trp 59, Tyr 67, Met 80, Phe 82 and the haem group were quantized with  $P = 10$ , which corresponds to a length scale of 0.1 Å. At this length scale, the potential, especially that of bond angles and dihedral angles which are the main variables in redox conformation switching, does not change significantly. Exploratory simulation of a single asparagine residue also indicated that  $P = 10$  is a good approximation. Solvent water was not included explicitly (except for three buried water molecules), but charge scaling and a distance-dependent dielectric constant were used; these have been shown to provide a reasonable approximation to equilibrium solvation effects<sup>1,17</sup>.

Table 1 lists the average deviations of various classes of atoms from the X-ray structure for the classical and quantum simulations. Both simulations provide the same qualitative observation: the dynamical deviations are smaller in the interior of the protein (within 12 Å of the protein centroid) than in the exterior. The shifts are smaller in the backbone than in the side chains. The shifts are larger for atoms further from the backbone. The haem group, which has a relatively rigid structure compared with the polypeptide and which is in the interior of the protein, shows a smaller shift than the rest of the protein<sup>17</sup>. The average deviations between the two simulations (listed in the third column) follow the same trend. But they are smaller than the deviations from the 1.5 Å X-ray structure, indicating a close and reassuring resemblance between the two simulations. Other

dynamic geometric features, such as atomic fluctuations in various regions in the protein, are also similar in the two simulations. In summary, the two simulations are both in closer agreement with X-ray structure than are most protein simulations, even those that explicitly include water<sup>1</sup>. More importantly, because the two simulations were started with the same atomic coordinates, the resulting structures are very similar and this provides a clear view of quantum effects.

There are significant differences in the classical and quantum distributions of many properties. Figure 1 shows the NH<sub>2</sub> dihedral angle distributions for Asn 52 from the two simulations. The width of the quantum distribution is larger, which is typical of quantum distributions. As in a harmonic oscillator<sup>12</sup>, the quantum effect should be more pronounced for larger ratios of  $(\hbar\omega)/(k_B T)$ . Figure 2 shows the Fe—N(His 18) bond length distributions. The quantum effect is more apparent because the bond stretching has higher frequency, whereas the quantum signature of low frequency modes in the protein is weak and can be buried in the noise in the simulation. Figure 3 shows the rotational distributions of one of the haem propionate groups. As the end of the side chain moves rather freely around the haem crevice and thus has relatively low frequency, it is not surprising to see that both simulations give similar wide distributions. The small quantum effect in the rotation of the propionate group is interesting in the context of redox conformation change: Williams and coworkers have proposed that hydrogen bonding with the propionate group is important in electron transfer reactions<sup>18</sup>. One would expect a small quantum contribution from the propionate rotation.

The work presented here shows the feasibility of quantum simulations of large biological molecules. By using this technique, we hope to examine the quantum effect of cytochrome *c* conformation switching and its influence on the electron transfer rate. For the nonadiabatic case this procedure can be easily extended for calculating the quantum nuclear contribution<sup>19</sup>. Such simulations can also be used to study the kinetics of proton transfer and certain conformational transitions, for example NH<sub>2</sub> rotation, where quantum dispersion and tunneling effects are expected to be significant.

This work was supported in part by NSF, the Robert A. Welch Foundation, the Texas Advanced Research Program, and the National Center for Supercomputing Applications at University of Illinois at Urbana-Champaign. J.A.M. is the recipient of the 1987 Hitchings Award from the Burroughs Wellcome Fund. P.G.W. was supported in part by the J.S. Guggenheim Foundation. The work described here was initiated while J.A.M. and P.G.W. were at the Institute for Theoretical Physics, University of California, which is supported by NSF and supplemented by funds from NASA.

**Table 1** Average deviations from the 1.5 Å X-ray structure for various classes of atoms

| Atom class            | Classical | Quantum | Classical-quantum |
|-----------------------|-----------|---------|-------------------|
| All atoms             | 1.71      | 1.71    | 0.63              |
| Interior              | 1.26      | 1.26    | 0.42              |
| Exterior              | 2.11      | 2.12    | 0.81              |
| Non-haem              | 1.76      | 1.76    | 0.65              |
| Haem                  | 0.53      | 0.62    | 0.28              |
| Backbone              | 1.27      | 1.32    | 0.49              |
| Side chains           | 1.87      | 1.87    | 0.68              |
| α-Carbon              | 1.30      | 1.36    | 0.50              |
| β-Carbon              | 1.39      | 1.42    | 0.53              |
| γ-Atoms               | 1.69      | 1.73    | 0.58              |
| δ-Atoms               | 2.06      | 2.05    | 0.72              |
| ε-Atoms               | 2.42      | 2.36    | 0.97              |
| Atoms quantized in QD | 0.81      | 0.83    | 0.36              |

All deviations are in Å. The average deviation of an atom is the distance of the average position of the atom in the simulation relative to the position of the same atom in the X-ray structure.

Received March 14; accepted 7 July 1988.

- McCammon, J. A. & Harvey, S. C. *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press, 1987).
- Frauenfelder, H. & Wolynes, P. G. *Science* **229**, 337–345 (1985).
- Tunneling in Biological Systems* (eds Chance, B., et al.) (Academic, New York, 1979).
- DeVault, D. *Quantum-mechanical Tunneling in Biological Systems* 2nd edn (Cambridge University Press 1984).
- Mayo, S. L., Ellis, W. R. Jr., Crutchley, R. J. & Gray, H. B. *Science* **233**, 948–952 (1986).
- Kihara, T. & McCray, J. A. *Biochim. biophys. Acta* **292**, 297–309 (1973).
- Albery, W. J. & Knowles, J. R. *J. Am. chem. Soc.* **99**, 637–638 (1977).
- Goldanskii, V. I., Krupnyanskii, Yu. F. & Flerov, V. N. *Doklady Biophysics* **272**, 209–212 (1984).
- Ringe, D. & Petsko, G. A. *Prog. Biophys. molec. Biol.* **45**, 197–235 (1985).
- Parak, F., Frolov, E. N., Mössbauer, R. L. & Goldanskii, V. I. *J. molec. Biol.* **145**, 825–833 (1981).
- Friedman, H. L. *A Course in Statistical Mechanics* (Prentice-Hall, Englewood Cliffs, 1985).
- Feynman, R. P. *Statistical Mechanics* (Benjamin, London, 1972).
- Chandler, D. & Wolynes, P. G. *J. chem. Phys.* **74**, 4078–4095 (1981).
- Berne, B. J. & Thirumalai, D. A. *Rev. phys. Chem.* **37**, 401–424 (1986).
- van Gunsteren, W. F., Berendsen, H. J. C., Hermans, J., Hol, W. G. J. & Postma, J. P. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4315–4319 (1983).
- Takano, T. & Dickerson, R. E. *J. molec. Biol.* **153**, 79–94 (1981).
- Northrup, S. H., Pear, M. R., Morgan, J. D., McCammon, J. A. & Karplus, M. *J. molec. Biol.* **153**, 1087–1109 (1981).
- Moore, G. R. et al. *Faraday Discuss. Chem. Soc.* **74**, 311–329 (1982).
- Wolynes, P. G. *J. chem. Phys.* **87**, 6559–6561 (1987).